Assessing the Validity of Latent Variables: The Cases of Social Values and European Identity

Thesis (cumulative)
Presented to the Faculty of Arts and Social Sciences of the University of Zurich
for the Degree of Doctor of Philosophy

by **Georg Datler**

Accepted in the Fall Term 2014
on the Recommendation of the Doctoral Committee:

Prof. Dr. Eldad Davidov (main advisor)
Prof. Dr. Michael Braun

Zurich, November 2015

Contents

Synopsis **5**

Paper 1

Testing the Invariance of Values in the Benelux Countries with the European Social Survey: Accounting for Ordinality **49**

Paper 2

Two theories on the test bench: Internal and external validity of the theories of Ronald Inglehart and Shalom Schwartz **75**

Paper 3

The Concept of European Identity: Overused and Underspecified? **121**

Paper 4

European identity as a safeguard against xenophobia? A differentiated view based on identity content **133**

Bibliography **163**

Curriculum Vitae ***179***

Synopsis

Contents of the Synopsis

Preface **9**
How to read this dissertation *10*
How this dissertation was written *11*

Prologue **12**
Where we are in the realm of latent variables *13*

1. The confusion with latent variables **15**
(My) confusion with latent variables *15*
What a latent variable is not *16*
Latent variable talk *17*
Proposing a framework *18*
What is (so problematic about) a construct? *20*

2. A detour to the philosophy of science: What is a latent variable? **21**

3. The issue of validity **24**
How many "validities"? *24*
The unified view: Messick *25*
The realist view: Borsboom *26*
In favor of realism II *27*
Construct validity—"an idea born dead"? *28*
The problem of quantitative measurement *30*
So what? *30*

4. Turning to the social world **33**
Different kinds *33*
Social science as re-construction *34*

5. Two case studies: social values and European identity **36**
CASE I: social values *36*
CASE II: European identity *40*
Conclusions: Social Science as Re-construction and Quantitative Methods *42*

Epilogue **43**

References **44**

Preface

A synopsis should be a glimpse of what the dissertation is about, i.e. it has the purpose to orient and summarize. A synopsis should also reflect and conclude. I have largely shifted the orienting function to the two short "how" sections in the preface. When writing the main part of the synopsis, I was attracted to reflect, to criticize, and to raise my voice. I chose to follow this impetus. Charles Tilly once wrote in the preface of his marvelous, little book, "Big Structures, Large Processes, Huge Comparisons":

> *Why do other people's books behave like docile marionettes? Mine keep playing Pinocchio. They take on characters of their own and resist correction. This one, for instance. When I sat down to write it, the book was supposed to end up mild-mannered, studious, and balanced: an even-handed survey of various ways to approach large comparisons of social structures and processes. Somehow it materialized as a bit of a bully. It struts around with a confident, pugnacious air. Yet behind the bravado hides a lazy, indecisive, pusillanimous weakling, with sticks for legs. My little book often makes accusations without naming names , avoids fights one might have expected it to welcome, and fails to specify when, where, or how alleged misdeeds occurred. Incorrigible! Yet, for all its faults, I love the little rapscallion.*
>
> (Tilly 1984, viii)

I couldn't have said it any better. I am not Charles Tilly but a PhD student, therefore I try to be balanced, respectful, and provide references as support for my arguments. There are many differences between Tilly's text and mine—quality, of course, and the topic. My little rapscallion is about latent variables and validity.

How to read this dissertation

The volume contains five parts—four papers and a synopsis. Every part can be read and is intended to be understandable on its own. This should especially be true of the four research papers; choose by interest. The first two papers are about social values, the other two are about European identity.Paper 1 “Testing the Invariance of Values in the Benelux Countries with the European Social Survey: Accounting for Ordinality”, is on the application of latent variable modeling to categorical data and assessing invariance of measurement across context—hopefully these clauses make clear that its emphasis is on applied statistics. If interested in substantive research on social values, Paper 2, “Two theories on the test bench: Internal and external validity of the theories of Ronald Inglehart and Shalom Schwartz”, is a better choice. Paper 3, “The Concept of European Identity: Overused and Underspecified”, is a critical examination of the concept of European identity, its uses, its definitions, and its problems. Therefore, it is a good choice if you want to get an overview of the debate on European identity. However, this overview might also leave you with confusion because the concept of European identity I regard as most tenable is but suggested in Paper 3 and more fully developed in Paper 4. Paper 4, “European Identity as a safeguard against xenophobia? A differentiated view based on identity content”, explicates what I hold to be a valid theoretical concept of European identity and applies this conceptualization in empirical analyses.

The synopsis should be the starting point if, and only if, you are interested in methodology. A bit of curiosity for philosophy of science and some tolerance to disillusionment would help as well. The synopsis can be read on its own; especially the first part of it. It is a detour to the history of quantitative social science and two of its central concepts, latent variables and validity. The second part of the synopsis is about social values and European identity as two cases of latent variables. It makes explicit, what is—hopefully at least to some part—applied in the four research papers.

How this dissertation was written

The synopsis was written last, which is not unusual. But in some way it did not behave as a “normal” synopsis should do, as a summary. Rather it dragged me deep into methodology and philosophy. This has two consequences—both of them I see as positive.

Some of what I have written in the research papers now seems problematic to me; which indicates that that I have learned something. The conclusions in the synopsis are more critical than in the original papers, although to the best of my knowledge nothing is utterly wrong in the papers. However, underlying issue stay latent in the research papers. The synopsis brings them to the fore.

The worst thing one can do with words is to surrender to them.
(George Orwell, Politics and the English Language, 1946)

Prologue

It is easy to name what this synopsis is about; it is about latent variables and validity. The terms variable, latent variable, and validity are part of the vernacular of empirical social science. Everybody uses them. Everybody knows what they mean. Not so; at least to me. I had to find out that I did not know the details, which is—as is well know—where the devil is.

Imagine my dissertation would have been a hike. It is always good to have a map with you when going for a hike. Actually, my map was a whole collection of maps, entitled "construct validity theory". It was a bit heavy, but I felt confident to have it with me on my journey. In paper 1 and paper 2, dealing with social values, I took a route that is sign-posted and actually I never had to use the map to find my way. Afterwards, dealing with European identity in paper 3 and paper 4, I felt ready to go into the open country, off the beaten tracks. I had my "construct validity theory" map with me. It did not take long and I wanted to use the map to figure out where to go. I stood in front of a forest. But there was no forest on the map. "Maybe a small mistake in the map", I thought and went on. After a while I was at the foot of a mountain; but no mountain on the map. I could not figure out where I was. I got lost and panicked, "Am I too stupid to read a map!?" But finally, I was sure, "The map is flawed!"

The mainstream view of construct validity did its job relatively well for what we wanted to do in the studies on social values. When turning to work on European identity, I got into severe trouble— some meandering had to do with myself, but some erring seemed to have to do with the methodological map itself. How hard I tried, I could not make standard conceptions of latent variables and validity work for European identity. I always had the feeling that I would force my thinking on the social reality. To be clear, good research always does some force to social reality; it reduces complexity. But in the case of European identity, the gold standards of measurement which I knew by then simply did not fit. I would say that there are good reasons why. Writing this introductory chapter was a welcome opportunity to put them down.

Where we are in the realm of latent variables

Before starting this endeavor it is wise to locate where we are. We are in the realm of latent variables; but where exactly? Skrondal and Rabe-Hesketh (2004, 1), in their seminal, unified treatment of modern statistics, start out with the observation that latent variables are omnipresent. "Omnipresent?" we could ask, "They are latent, i.e. rather hidden, aren't they?" Of course, this is a misreading of Skrondal and Rabe-Hesketh; they want to say that latent variables are used for a wide range of purposes. But there is something to the naïve reading of omnipresence—it triggers the question what the nature of latent variables is. We will see later that the answer depend on metaphysics.

I hope that reading "metaphysics" already triggered a reaction—be it curiosity, doubt, or shaking the head. While there are good reasons to classify some metaphysics as ridiculous, we will see that we need an account of ontology to answer what a latent variable is. Ontology is the branch of philosophy that asks about the fundamental nature of the world and being. Its non-metaphysic sibling is epistemology, the branch of philosophy that asks about the possibilities, ways, and forms of knowledge. The ontology-epistemology pair will turn up again and again. If unfamiliar, a good way to follow is to keep in mind that ontology refers to the question "*what is real?*" and epistemology refers to the question "*How can we know?*". These two questions also make clear immediately that answering epistemological questions presupposes—if implicit—an ontological position. Statistics is not as philosophically innocent as it may seem.

Before moving on, we take "omnipresence" as Skrondal and Rabe-Hesketh intended and survey the usages of latent variables. This will help to single out which usage I focus on. Amongst the most important usages of latent variables are:

(1) Decomposition of variances and covariances

(2) Modeling unobserved heterogeneity

(3) Missing data and counterfactuals

(4) Hypothetical constructs, latent responses

Usages (1), (2), and (3) are technical in the sense that they propose latent variables as a way to handle problems in the observed data. They are not in the focus of my attention (excellent treatments are Bollen 1989, 2002). The fourth usage of latent variables is not technical but theoretical. Skrondal and Rabe-Hesketh (2004), in line with many others, use the term "hypothetical construct" to refer to latent variables as concepts and the term

“latent response” to refer to (continuous) response variables underlying observed (categorical) variables. I will take issue with the term “construct” later. For now, it is only important to note that latent variables denote on the one hand technical stuff and on the other hand theoretical stuff.

The focus of the following is on theoretical stuff, i.e. on latent variable as concepts. I will not use equations or formal notion; others are better at this. Yet, the claim is to relate theoretical reflection to scientific practice, part of which is statistical calculus. In this sense the text is more about quantitative methodology than about specific quantitative methods.

The text proceeds in the following steps, some of which may seem unnecessary—they are indeed, if you are really into methodology. Speaking of my own experience, there is potential for fallacy here. I assumed that I was really into methodology because I knew the methods everybody uses, and some more. However, I had to realize that running a structural equation model is something different to methodological maturity. As some of the following is “not mainstream”, it might well be that you feel comfortable in your “mainstream” environment, without having an idea what issues lurk behind the surface. A good way to decide whether to read the text from the beginning to the end is to the look at the opening example of the next chapter, the three different interpretation of the factor model. If it is obvious to you which interpretations go with which ontological and epistemic assumptions, go ahead to any chapter you find useful. If not, take the stepwise, narrative approach with me. Here is the road map:

(1) The confusion with latent variables

I start out with describing my confusion; of course hoping to convince you that there is confusion in the scientific field and that we should proceed cautiously. I will first approach latent variables in the negative, i.e. in terms what they are not. Afterwards we will have a glance at latent variable talk and realize that it has something of a Babylonian confusion of terms. In order not to surrender to these words, I introduce a three level framework to talk about latent variables. Using this framework, the issues of latent variables can be articulated clearly. Especially, we will have a first try at disenchanting the term “construct”. “Construct” seems to be at the centre of the problem. “Construct” seems to do a trick; to be magical. This is bad metaphysics.

(2) A detour to philosophy

We need some good metaphysics. In order to fully understand what the trick of "construct" is, it is necessary to spell out two philosophical positions on what latent variables are, constructivism and realism. I will defend a realist position.

(3) The issue of validity

We will see that a realist notion of latent variables gets the issue of validity different to the tradition of construct validity. At this point, I will bring all the strings together in a criticism of construct validity, with a side step to the related issue of measurement. Of course, crushing construct validity, leads to the question what to do instead. I will argue that being less afraid of theoretical work and being less detached from social reality are helpful.

(4) Turning to the social world

Therefore, in the fourth part, the synopsis turns to the social world and claims that human kinds are different from natural kinds.

(5) Two case studies: Social values and European identity

Finally, I will be able to explain why sociological concepts such as European identity are a hart nut for quantitative science, and how I tried to crack it. Having read the other chapter, I hope the reader will be in a position to judge how badly I failed by my own standards.

1. The confusion with latent variables

(My) confusion with latent variables

Let me use an example to show how confused I am. I run a factor analysis on three variables x1, x2, x3 which results in one factor X. Now, I try to interpret the result and come up with three options:

(1) "I have attempted to measure construct X."

(2) "I have constructed a hypothetical variable out of x1, x2, x3."

(3) "I have summarized the data into a composite made of the shared variance of x1, x2, x3."

All three answers sound more or less appropriate to me. One might chip in that the answer depends on whether I had run an exploratory or a confirmatory factor analyses. This is the right track, though the wrong angle. The angle to start with is theoretical and conceptual, i.e. outside of the data (analyses). And maybe, with a more basic topic: The semantic relation of the terms variable and latent variable.

What a latent variable is not

Latent variables are commonly described in terms of what they are not; they are not manifest; they are not observable or not directly observed. This is helpful as long as we know, what variables are. Let us pretend not to know and start from two simple propositions about variables:

(1) A variable is something that varies.
Explanation: This excludes constants.

(2) A variable can be coded in numbers.
Explanation: A variable has different categories (states) or its quantity (amount) can vary. More precisely, not the variable is coded but the position of object on the variable.

Age is an example of a variable conventionally treated as manifest. Taking the two propositions above, age is the amount of time passed since birth coded in years. This is unproblematic. However this sentence accomplishes a lot:

(1)	Age gets a definition.	CONCEPT
(2)	The definition refers to a distinct aspect of reality;	REFERENT
(3)	We can observe this aspect of reality and there exists a metric to code the observation.	MEASUREMENT

All this is meant when we say that we successfully measure a variable (Michell 2013). The first two accomplishments concern ontology (What is real?); the third accomplishment concern epistemology (How we can know?).

Latent variables are often portrayed against observed variables—but what often goes unnoticed it that both are variables. All variables need a definition, a referent in reality (more about this later), and pose the problem of how to measure them. In the next section I will argue that much of latent variable talk amounts to a Babylonian confusion that obscures rather than clarifies these issues.

I have already smuggled in an important issue in the first sentences of this chapter. I said that latent variables are "not observable or not directly observed". These two propositions make a difference. "Observable" vs. "not observable" is a categorical statement on the nature of variables. This is an ontological distinction. If correct, there are two kinds of variables: Variables we can observe and variables we cannot observe. Of course the question then is what latent variables are. Can they be said to exist? The second proposition, "not directly observed", only states that a variable is latent when we did not or could not observe it. This is an epistemological explanation why some variables in our models are latent and some are manifest. In this view, both observed and latent variables can be real or not.

In the course of the next sections we will see that the latent variable framework starts out from the epistemic issue of accessibility but seems to solve ontological, conceptual and measurement issues passing by. Do these functions do us a favor? To judge, we need to decipher latent variable talk.

Latent variable talk

The quantitative approach entered the stage roughly a hundred years ago with the forceful premise to make social science more scientific. Scientific—at these times rather unambiguously—meant evidence based, objective, in search of general laws, and eventually truth seeking. Latent variable modeling is one of its most prominent offspring.

Ironically, what began with the intention to make the study of the social and the psychological more like the perceived ideal of natural science, resulted in a plethora of concepts, in an uneasy amalgam of ideas, and in confusion, which found its smallest denominator in the employment of statistical machinery. We all work with some kind of statistical models that entail factors, classes, or components, but we rarely admit that we are confused. The reading which is most in favor of us is that we are simply used to latent variable talk.

A less favorable reading is that we have surrendered to latent variable talk—maybe up the degree that we do not see the issues anymore. Here is a reminder on the confusion in latent variable talk. The term "latent variable" is frequently used synonymous with the terms "construct", "factor", "concept", "trait", "unobservable", and "hypothetical construct" to name but a few. These terms that should tell us what a latent variable is, have themselves acquired different meanings and connotations. A few examples will

suffice (for a critical overview, see Maraun and Gabriel 2013): Is a latent variable understood as a factor the same as the factor in a statistical factor model? Is a hypothetical construct hypothetical in the sense that it is hypothesized to exist or in the sense that it is constructed out of the data. Are latent variables as concepts part of the real world or just linguistic symbols?

This play on words could go and on. The examples I gave, have a common structure: They take a term and use it at different levels of abstractions. Take the second example, it uses the term "hypothetical construct" first as a theoretical term and then as term that describes data manipulation. Following Rigdon (2012, 2013) latent variable talk is about three levels of abstraction: concepts, proxies, and observed variables.

Proposing a framework

Figure 1 proposes a framework to talk about latent variables. In the three levels of abstraction it follows Rigdon (2012, 2013); the naming of the levels and object is modified and the concept/referent distinction are added, which is in line with the thinking of Rigdon (personal communication, July 23rd, 2014).

Figure 1 The concept proxy framework

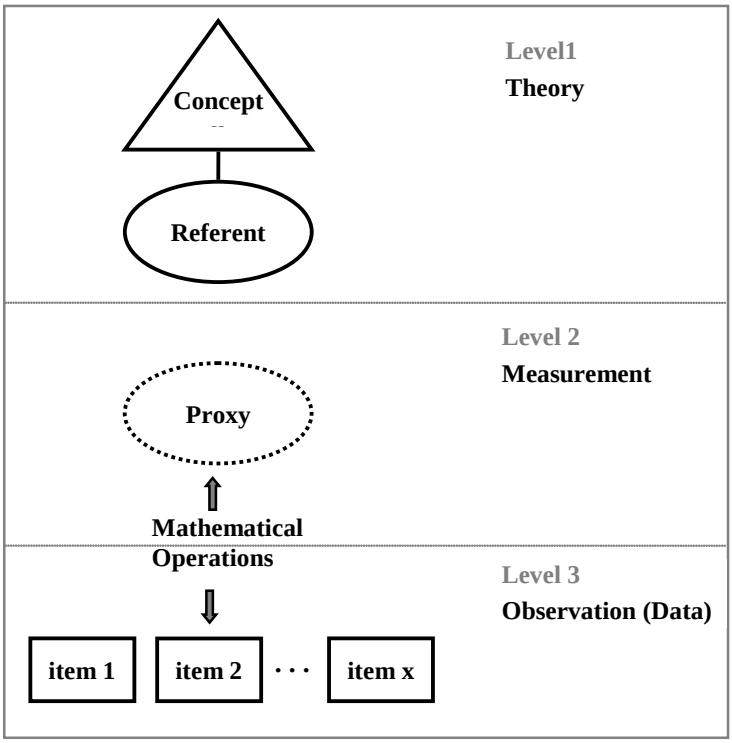


Figure 1 is about latent variables; the term "latent variable", however, is absent. "Construct" is missing as well. Apparently, I do not appreciate them. I will try to

convince the reader that the proposed alternatives do a better job to describe what we do when we work with latent variables; and that this framework allows to understand why much of latent variable talk rather obscures than clarifies the issues we (should) encounter.

At the most abstract level (level 1) is the theoretical concept. A further distinction at the theoretical level is the so-called *concept/referent distinction* (Maraun and Peters 2005). The concept as such defines a class of entities; a concept denotes. The denoted entities are its referents. This may sound like hairsplitting. I will try an example: Level 1 looks light a tree, right? In a nutshell, the concept/referent distinction says that you can't climb the word "tree". The word "tree" is a concept in language, a symbol. Its referents are out in the world; these trunks in your garden you can climb. Why is this important for methodology? A simple analogy might touch the point: If you do not know what you are looking for (concept), it is hard to find (referent). You cannot find out what you were looking for by grabbing something at the wayside and proclaiming "Ah, that's what I was looking for". In terms of latent variables, the concept/referent distinction means that we cannot fix the definition and the meaning of a latent variable by the data alone. Or put differently, we are urged to explicitly define our concepts; for there must have been an implicit understanding of what we were looking for in the first place when we collected data. Stated more plainly, factor analyzing a set of items will not substitute for theoretical definition.

At the lowest level of abstraction (level 3) is the observational information we have at hand. Most commonly, these data are called observed variables. This term is unproblematic and I will use it as well, though I prefer to call the data a set of items. "Item" underscores that it is on open question whether are not these data entries have the properties of a variable and whether or not they relate to a latent variable.

At the intermediate level (level 2) there are the latent variables in their technical sense. They are supposed to stand in as proxies for the concepts (more precisely, the referents of the concepts).[1] All proxies are the result of mathematical operations on the items. Factor methods yield factors proxies; component based methods yield composite proxies. The

[1] As the concept/referent distinction is not widely used in the social sciences, I will in most occasions use the term concept and add a qualification when I mean its referent. "Referent" alone might not be too unfamiliar. From a methodological standpoint, this distinction is important—its absence opens the door to the nominalist fallacy.

promise of latent variable modeling is that specific mathematical operations (e.g., the common factor model, the Rasch model) results in the successful representation of theoretical concepts. Measurement is a special kind of representation; we will touch upon this issue later.

What is (so problematic about) a construct?

Construct is one word for two different things. First, it can mean a theoretical term. A theoretical term is a linguistic, symbolic entity that serves as a placeholder in theories. This is what figure 1 calls a concept. Second, it can mean the properties that we think play a role in reality and of which we would like to obtain measures. This is what figure 1 calls referents. Polysemy may be annoying but is the normal case in language, no problem as long as one would clearly recognize which of the two meanings is intended. This is not the case with the usage of construct. One regularly sees properties that can be attributed to construct in one sense being transported to construct in the other sense (Borsboom et al. 2009). It seems that the function of the term construct is exactly this blurring of distinctions. The blurring of the concept/referent distinction concerns the theoretical level. An even more severe "concept equating" (Maraun and Gabriel 2013) takes place across the levels of theory and measurement: The term "construct" is used to denote both the theoretical concepts at the most abstract level and the proxies at the intermediate level (Rigdon 2012).

This is confusing; and it has the smell of cheating. This is not to say that everybody who uses the term "construct" intends to cheat, nor that those who came up with the idea intended to cheat. But it is to say that a whole lot of problems are connected to "construct". A first example: The blurring of distinctions triggers the belief that the empirical representations are per se equivalent to the theoretical concepts (Maraun and Gabriel 2013). Despite the warnings of Cliff (1983) and others about the *nominalist fallacy*—thinking that something is identical with whatever label has been applied to it—researchers in the factor-analytic tradition have been particularly inclined to assume that the common factors derived in their statistical models are equivalent to the conceptual variable labels assigned to them.

We will take up this thread again in the section on validity. The reason for yet another adjournment in the story is that my argument rests on an ontological commitment to realism about latent variables. I have already made use of a realist perspective when I

echoed the warning against the nominalist fallacy and when I made an issue out of the concept/referent distinction. A detour to philosophy will present realism and its main rival, constructivism, and spell out the reason why only realism about latent variables is tenable. The three levels of abstraction—concepts/referents, proxies, items— will serve again as an orientation.

2. A detour to the philosophy of science: What is a latent variable?

There are a range of formal and empirical definitions of latent variables (Bollen 2002, Borsboom et al. 2003). However, statistical theory only tells us how parameters that relate the latent variable to the data could be estimated, if the data were generated under the model in question. The *if* in the preceding sentence is decisive. It points out that latent variable theory includes assumptions of how the world works. How we structure our models should have something to do with how the world works. Latent variable modeling is not neutral on ontological issues (Borsboom et al. 2003). In principle, two extreme views can be contrasted: latent variables as fiction (constructivism) and latent variables as real (realism). Constructivism and realism are themselves heterogeneous schools of thought. Though compatible with many constructivist and realist accounts my usage is restricted to the object of latent variables, i.e. best described as a constructivism or realism about latent variables (for constructivism and realism about the social, see section 4).

Constructivism: Latent variables as fiction

The estimation of latent variable models is, of course, dependent on the input of observable information. The output of latent variable models is all there is about latent variables. There is no latent variable unless we compute it. In this view, latent variables are pure constructions; they are not real. Therefore, latent variables are a numerical trick—albeit a potentially useful trick. Latent variables can serve as elegant heuristic summaries of a bunch of observed variables and as taxonomic tools to establish classifications. But they cannot stand in for theoretical concepts, simply because there are no theoretical concepts that are real (more precisely, there are no theoretical concepts that have empirical referents).

Paradoxically, a similar attitude towards latent variables developed within the school of logical empiricism ("positivism"), though for different reasons. The early positivists gave

priority to characteristics that are directly observable. Eventually, they realized the need to include in their theories unobserved concepts and came up with operationalism. Operationalism holds that the meaning of concept is synonymous with the set of operations used to measure it (Bridgman 1927). Operationalism allowed the positivists to uphold the primacy of the observable, but turned out to create another problem. If the measures define what the latent variable is, there are as many latent variables as sets of measures. There is no way to decide which measures and which operations should be used, apart from an instrumentalist point of view (Block 1976): Latent variables are a means to an end, be it prediction or data reduction.
Operational definitions cling to the observable; but it is difficult to defend what is an observation and thus real and what is a unobservable and thus a construction (In philosophical jargon: It grants ontological significance to a sharp distinction between observation and theory and claims that only manifest, or observable entities exist, when observability is really a matter of degree (Haig 2013). A simple example is the magnifying class. Through a magnifying glass you see small things you did not see before. It seems plausible to admit that the small things have existed before. Regarding latent variables as fiction has some shortcomings and some cumbersome consequences.

Operationalism differs from philosophies of constructivism in that it (uncritically) accepts the realness of observations. For our subject matter, however, this does not make a difference. Whereas the operationalist accepts observed variables at level 1, the constructionist may or may not do so; both of them, however, do not allow real entities at level 3. When there are no real concepts (or more precisely: no real referents of concepts), it makes no sense to talk about proxies at level 2. Data constructions is all there is.

Latent variables as real

A realist maintains one or all of the following propositions: (1) theories can be true or false, (2) Theoretical entities, at least some them, exist, (3) Theoretical entities are causally responsible for observed phenomena (Devitt 1991).

Borsboom et al. (2003) argue that realism is the only ontological position that is compatible with latent variable modeling. Both factor models and item response theory (IRT) models treat the latent variable as the source of the observed variation in the indicators. The latent variable causally explains the observed pattern. Only if the assumed concept exists, the latent variable proxy can stand in for it in the model.

This argument seems a bit like begging the question because it starts from the observation how factor models view the world and not from the observation how the world is and which models are appropriate as a result of it. After all, the claim that latent entities have to be real to make factor models meaningful, serves a as reminder to the practitioners of latent variable modeling. It is not factor analysis that constructs meaningful factors. On the contrary, factor analyses are only warranted if the belief is justified that meaningful latent entities exist.

Let us reframe the realist position in terms of the three levels of abstractions. Also in the realist account, concepts remain theoretical, but they do refer to latent entities that exist (their referents). The latent variables in the statistical models are proxies for these latent entities; therefore proxies can be modeled as the common cause of the observed items.

In favour of realism I

I feel I should start with what is not my intention: Realism about latent variables is no carte blanche for latent variable models of the common-factor type. Rather, realism spells out the condition under which common-factor type models are tenable: A latent variable has to exist. Computing a latent variable model does not make a latent variable real. If a latent variable exists, it is appropriate to compute a model that could works as a proxy for it. That said, if a latent variable does not exist, pull the plug.[2]

That said, I can safely turn to batter the view that latent variables are fictions. However, again a qualification: More precisely, I criticize the constructivist usage that wants to make realist interpretations. Intentional constructivist usage as pure heuristic device is a viable option.

Constructivist philosophy is skeptical about science; albeit very fond of science, operationalism left us with a similar result. If the meaning of the concept is defined by the set of items that should measure it, every different set of item gives rise to a new concept. Already back in the 1980s Duncan (1984, 227) lamented the thousands of scales that

[2] Of course, in reality we do not know for sure if a latent variable exists or not. But again, this is no carte blanche to claim any latent variable. A successful claim of a latent variable must give a plausible account how this variable operates in the world. This is more demanding than it may seem; psychologists, for example, ponder over the realness of "intelligence" since the days of Charles Spearman (Borsboom 2005; Gould 1996).

sociologists had produced and their attitude that all would measure something and their attitude that they all would measure something different.

We should add that if you really regard latent variables as fictions, measurement is an oxymoron — you cannot measure something that is not real. Clearly, the early operationalists were not fully aware of the consequences. Sadly, the notion of operational definition still persists.

Much of the confusion around constructs and latent variables stems from an uneasy amalgam of realist intentions and constructivist practices. Construct validation theory is no exception. In its simplest form the notion of validity has something with the issue whether a test measures what it should measure. We will see that the aspect of measurement is not at the forefront of the current mainstream of construct validity theory. However, try to think about validity without thinking about measurement. I bet it won't work. When we agree that measurement is an issue, it does not make sense to adhere to constructivist practice and to pretend to measure something, let alone to claim valid measurement. Unfortunately, something like this seems to go on in construct validity theory.

3. The issue of validity

How many "validities"?

Standard textbooks on research methods (e.g., Crano and Brewer 2002, Schnell et al. 2008) give the impression that there were several types of validity. It is more useful to frame these approaches as different strategies of validation. Validity is a single property; validation is the activity to look for validity evidence.

Table 2 shows different approaches to validation. It is apparent that the main tool for quantitative validation studies is correlational analysis (which includes factor analysis). This is a rough classification, and though the terms are still in use, it is outdated. Tow developments of the last decades deserve more attention: the unified view and the realist view.

Table 2 Types of validity evidence

type	evidence	how to
content	Measures represent the theoretical domain of interest	reasoning; intersubjective agreement on meaning
criterion	Association of the measures under scrutiny with other, well-respected measures Original: criterion = measure of the very same concept Later: criteria = different concepts, "concurrent validity", "divergent validity", "predictive validity".	correlate, correlate, correlate
construct	The relations of construct X with other constructs should conform to the theoretical expectations (the idea of the nomological network)	deduce and then correlate, correlate, correlate

The unified view: Messick

In the legacy of Cronbach and Meehl (1955), a unified view on validity developed that subsumed all approaches to validation under the umbrella of construct validity (for detailed historical accounts see, Sireci 2009, Newton 2012).

Other labels for this approach are "mainstream" and "consensus"—the reason being that it was encoded in an important guidance document, the Standards for Educational and Psychological Testing, published jointly by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (1999, 2nd edition).[3]

Messick, a key author in the unified view, holds that "validity is an integrated evaluative *judgment* of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment" (Messick 1989, 13, italics in the original). Let us go slowly through the definition. Messick proposes that validity is a judgment not a property. To be judged is the use of the test, inferences and actions, not the test itself. Validity evidences are all empirical observations and theoretical rationales that speak about the adequacy and

[3] A revised, 3th edition is expected to be published in 2015.

appropriateness of test use. This definition is broad enough, to give semantic concerns and correlation strategies its say in validation research. The emphasis is, however on pragmatic the pragmatic aspect of validity, i.e. what is to be done with the test scores and how testing is done (Kane 2006; Zumbo 2009). This includes usage, explanation of the testing process, and ethical issues.

For its all-encompassing nature the unified view has been criticized as a "laundry list" of validity evidence—in Messick's account it is not clear what is sufficient for validity and where, except everywhere, to look for it. This disorients researchers who keep on computing correlations (Hood 2008). Everything that is in the unified view of validity is important—intended use, consequences, ethics—but is it reasonable to pack all into validity?

The realist view: Borsboom

The answer of Borboom and colleagues would be: No, it is not reasonable to put all desirable things into the definition of validity. In their view, the unified view at best hides the key ingredient, at worst it is lacking: Successful measurement. While the mainstream view of construct validity, treats validity as an evaluation of *test use* or *test interpretation*, Borsboom insists that validity is a property of *the test* (Borsboom et al. 2009). In one sense, Borsboom is close to the classical view of validity, before the reign of construct validity, which holds that validity is the degree to which "a test measures what it is supposed to measure" (Garrett, 1937, 324). Validity is an attribute of a measurement instrument that tells whether the measurement instrument is sensitive to the variation in the targeted attribute. Borsboom et al. (2004, 1061) propose "a simple conception of test validity: A test is valid for measuring an attribute if (a) the attribute exists and (b) variations in the attribute causally produce variation in the measurement outcomes."

But they add what they call latent variable theory (Borsboom et al. 2003). Besides statistical models, latent variable theory contains a realist ontology. If validity is about measurement, there has to be something to measure in the first place. Latent variables need to exist, to judge how well a test measures them. If the test scores shall represent the position on a latent variable, the measures need to reflect the latent variable. Following Borsboom the best candidate for this reflection is causality. Latent variable theory assumes that latent variables causally produce the answers to the items. In this view, reflective models are a good candidate for measurement, formative models are not.

Reflective models, such as factor and item response theory (IRT) models treat indicators as the effects of latent variables. Formative models put it the other way round: Indicators cause the latent variable (Bollen and Bauldry 2011, Bollen and Lennox 1991). In a realist perspective, formative models are composites of observed variables but not measures of latent variables (Lee et al. 2013).

In favor of realism II

A realist view on validity is not only clearer than the unified view but logically more stringent. The claim that a concept exists is a necessary condition for measurement to be possible—there has to be something to measure. Disregarding cases where wrong theories lead to correct conclusions, validity understood as successful measurement is a necessary condition for appropriate test use. It seems odd to investigate test use in the sense of construct validity without a prior examination of ontology and measurement.

Validity, then, is a function of truth and not of evidence (Borsboom et al. 2009). Yes, truth. *Truth* is one of the terms—among them probably *existence* and *reality* as well—social scientists have learned to avoid. If they are really afraid of them, they should stop talking of their latent variable models as measurement models. Validity, as the claim that a latent variable exists and causally produces variation in the observed items, is a claim about how reality is—In other words, a claim about truth. Truth claims, however, neither mean that science reveals the truth nor that science can discover the truth. Truth claims make it possible that science aims at truth, however badly we fail. Maybe a well liked statement about statistical models can do an illustrative job: "All models are wrong"—this is correct as an observation of modeling practice; not so as a general proposition. The whole notion of misspecification requires that a true model exists and that we judge how close our model comes to it (Borsboom 2005, 66). A side note to get some grips on our philosophical siblings: What I have just said is that the assumption "All models are wrong" is warranted in epistemological terms (All models I have seen are wrong) but not in ontological terms (No true model can exist.).)

Traditional construct validation seems to get the process backwards: Apply a test and then determine if you are measuring what you intend to measure. Though it offers plenty of suggestions how to do this, it seems not what we want to. Rather we ought to use knowledge of the concept and knowledge of the causal properties of the referents of the concept, to construct a measurement instrument. The challenge, of course, is to get the

knowledge of the concept. At least it should be clear, that statistical models cannot generate this knowledge. Either we are lucky and there is an elaborated theory of the concept, or we have to engage in theoretical work. *Theorizing* may be yet term that causes distress. Theories, we have learned are a good thing, but our job is to deduce from them and not to make them. Where do theories come from? To see how we could get so perplexed by this question, I will—once again—turn to the story of construct validity.

Construct validity—"an idea born dead"?

In a realist perspective, only models that assume the latent variable to be the common cause of the observed items qualify as measurement. The reverse, of course, is not true. Setting up a factor model, does neither make a latent variable real nor does it ensure successful measurement. Though put a bit too simplistic, this is what plagues the domain of construct validity. Critics have not been cautious when pointing at the deficiencies of construct validity and psychometrics in gerneral. There is some polemic at play when Michell (2000, 2008) denounces the practice of psychometricians to be pathological and Maraun (2007) speaks of latent variable models as "myths and confusions". We will turn to the history of methodology to understand how good intentions—to make social sciences more scientific—could bring about a state of affairs that triggers such offensive remarks.

Borsboom et al. (2009, 137) proclaimed the "end of construct validity", convinced that "the idea of construct validity was born dead". If so, it showed a remarkable strive to come alive and stay vibrant. Researchers routinely speak of the validity of their constructs; no textbook on research methods can do without a reference to construct validity. Of course, Borsboom et al. know this; what they mean is that the idea of construct validity was flawed from the very beginning. And, one may add, that it got even more problematic in its adolescence when the idea left philosophy of science and settled down in the milieu of applied science.

Enough metaphors, Let us jump into the story at the beginning of the 20th century, the heydays of logical positivism, a time when psychologists and sociologists wanted to make their fields more scientific (for more details see, Duncan 1984, Maraun 2007, Michell 2013). Logical positivism gives priority to directly observable phenomena, theoretical concepts can only be constructed out of observed phenomena. This implies that the meaning of the resulting constructs can only be due to its observed constituencies. Some

decades later, rebranding the program to “logical empiricism” Carnap (1936, 1937) still prioritized that constructs get their meaning from observable terms (via what he termed correspondence rules), but acknowledged that the meaning of constructs can never fully be reduced to observables (“surplus meaning”). Hempel (1952) described the relations as a “complex spatial network”. In Hempel’s analogy, theoretical concepts are the knots, definitions and hypotheses are threads connecting them and the entire network is thought of as anchored to observed concepts by strings, which correspond to rules of interpretation. We are now close to what became the standard view of social research. We only need make one more observation. Psychologists these days embraced positivism; but were interested in abstract phenomena, such as intelligence. Their way out was to employ “operational definitions” of concepts, which, following Bridgeman (1927), say that a construct is equal to the operations used to measure it. Operationalism confuses *what* should be measures with *how* it is measured (Michell 2009) and leads to an absurd proliferation of concepts; every new set of measures gives rise to a new concept. In this climate Cronbachs and Meehls famous work on the “nomological network” and “construct validity” can be seen as an attempt to rescue psychology from operationalism. Cronbach and Meehl insisted that “the meaning of theoretical constructs is set forth by stating the laws in which they occur” (1955, p. 294). There view of construct validity depends on laws that connect constructs with other constructs. The crucial point is that psychology and for that matter sociology had no laws, only alleged connections between constructs, and correlation coefficients between variables (Cronbach and Meehl allowed correlations and factor loadings to substitute for a network of laws, hoping that scientists would eventually be able to precisely define their constructs “when all of the laws involving it” were discovered (Cronbach and Meehl, 1955, 294)). This was 60 years ago.

Waiting for the laws to come, we went one with constructs as empty “promissory notes”(Messick, 1989, 23) for theoretical concepts. In the interim, operational definitions sufficed for the pragmatic purpose of anchoring concepts to observable criteria. What intended as promissory proved here to stay; the understanding of concepts is deferred indefinitely.”However, ignorance regarding the meaning of constructs is a significant obstacle if one wants to claim that tests measure them” (Michell 2013, 15).Yet, validity is exactly the premise that a test measures a concept. The notion of construct seems not very helpful to issues of validity; and for that matter we can also touch on another, related problem.

The problem of quantitative measurement

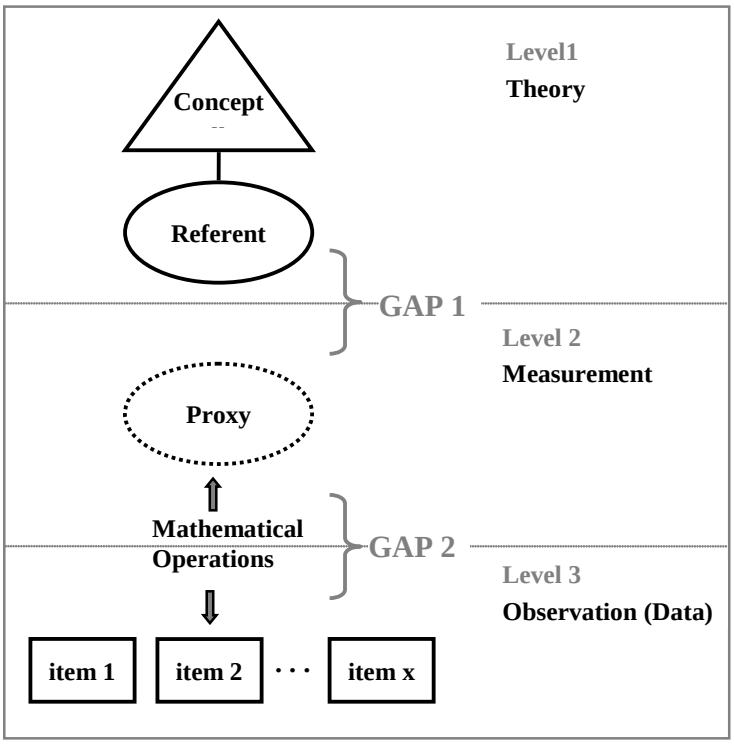
It is no coincidence that the definition of measurement most widely cited in the social sciences is operational. Stevens (1946) defined measurement as the assignment of numbers to objects according to rule. Again, this states *how* to measure but not *what* to measure. With some polemic Michel (2013) explains that only quantifiable properties can be measured. Measurement in the strict sense demands that the measured property has a continuous quantitative structure, i.e. in Stevens terminology is on an interval or scale). What we know as nominal or ordinal scales is not measurement in this sense—to use different terms, it is representation or assessment (Maraun and Peters 2005). However, it is important to note, that such categorical information is valuable, the normal case in social sciences, and maybe sufficient (Guttman 1971). But it is not measurement in the sense we can measure distances with a meter stick.

The crucial point is that the notion of construct and construct validity was used to wrongly pretend that the social sciences know how to measure. Though the coded observations are categorical, the latent construct does the trick: It is metric by definition (at least in factor model). Construct validation brings the trick to perfection. Test scores—whether factor scores or some other index does not matter here— are quantitative. If test scores and their relations are defining the construct, the construct must be quantitative. Using the three levels of abstraction, we once again see how construct (validation) theory obscures the level of the concept and the level of the proxie. A factor model makes the proxy quantitative, but not the concept.

So what?

I have argued in favor of a realist conception of latent variables and of validity as successful measurement. These two are linked. Messick (1989, 26) concisely expressed this point, “One must be an ontological realist in order to be an epistemological fallibilist”. You can only be wrong about things that can be right. Taking a constructivist route leads somewhere else: Validity then cannot be about measurement. For if there is no real latent variable, there is no true position on the latent variable, and you can never be wrong about it. Working with fictions you can only judge their usefulness, which is of course dependant on your intentions and on the task at hand.
Compared to the correlational dogma, realist psychometricians have been lean on practical advice (Hood 2008). Let us employ the three levels of abstraction one last time.

Figure 2 validity gaps



In a realist framework there are two gaps to worry about: Gap1 between concepts and proxies, Gap2 between proxies and observed items. Construct validity theory deals exclusively with Gap 2 between proxies and data because they collapse proxies and concepts into constructs. Validation in this sense means fitting a measurement model to the data. Such a model posits that latent variables are the causes of observed answers; model fit and the strengths of measurement parameters signal the degree of validity. The crux is, this strategy works fine, if and only if validity Gap 2 is already closed. One issue is that the researcher has to find the right measurement model; is it a factor model, an item response theory (IRT) model, IRT of the Rasch type, IRT of the Mokken type? It is a myths—though a persistent one—that model fit alone could help with this. There will always be several models that fit the data equally well, a fact known as the problem of equivalent models (Lee and Hershberger 1990, MacCallum et al. 1993). Especially in factor analyses the predicate "confirmatory" seems to invite wishful thinking. The warning of Cliff is not outdated still: "Even in what is called 'confirmatory' factor analysis, it is not the nature of the factors which is confirmed: the only thing which is confirmed is that the observed covariance matrix is not inconsistent with a certain pattern of parameters. It does not tell us what those parameters mean, and experience has shown that our belief that we do know what they mean is often ill-founded" (Cliff 1983, 122).

Little example? The well-known theory of personality which is marketed as “Big Five” might as well be “The classical Six”. The theory was “validated” (or developed) with exploratory factor analysis which reduced the data the five dimensions. Statisticians know that a factor model with k continuous latent variables is equivalent to a latent class model with k+1 classes (Bartholomew 1987) Had the researchers used latent class analyses, we would now talk about The Big Six types of personality instead of The Big Five personality dimensions (Borsboom et al. 2003). Psychometrics alone cannot help in this case. What psychometrics can do, is translate. Psychometrics can translate assumptions on the nature of a concept and assumptions how the concept is related to observed data into model constraints. In principle, psychometrics is well equipped to close Validity Gap 2 (Borsboom 2006), if it has enough knowledge at hand. Conceptual knowledge, however, does not appear from nowhere.

Validity Gap 2 is exactly concerned with the relation of theoretical concepts and latent variable proxies. I think it is this gap that deserves equal attention. This means theoretical work. Among many researchers in the hypothetico-deductive tradition, theory is something others do. In fact Popper fueled this view when we wrote that “The act of conceiving or inventing a theory, seems to me to neither call for logical analysis nor to be susceptible to it” Popper (1959, 31). The hypothetico-deductive approach to theory misconceives theory formulation as pure construction. Following Michell (2013) viable theory production is first and foremost reasoning from phenomena. To reason from phenomena, with the intention to formulate a theory that explains, requires first of all looking at the phenomena.

I would like to conclude this section on validity with a small example of looking at a phenomenon. Observing newspapers and feuilleton, I saw that one and the same commentator change his opinion on European integration several times. At the same time, I learned that a realist interpretation of a factor model fits well with so-called traits. According to Messick (1989, 15) a trait is “a relatively stable characteristic of a person—an attribute, enduring process, or disposition—which is consistently manifested to some degree when relevant, despite considerable variation in the range of settings and circumstances”. This observation made me question whether a factor model could measure European identity—*measure*?

4. Turning to the social world

I agree with the realist conception of latent variables as proposed by Borsboom and colleagues. Again, some hiking metaphor: Borsbooms vantage point was psychometric modeling and he arrived at the worldview that is entailed by the latent variable model. Though Borsboom is critical enough of dubious practices of latent variable modeling, his approach is a bit like begging the question. He starts out with the factor model and asks which kind of reality must exist to justify the factor model. I will start out with a realist notion of social reality, and see if I arrive at the factor model.

Different kinds

What I have written so far applies to stones, without doubt. Stones are the objects and gravitational force is the latent variable that makes them fall down. The methodological literature in the social sciences is full of examples from the natural sciences—nothing wrong about that; as long as one does not equate humans with stones and love with gravity. There are some differences: Stones do not care about the theories we have about them; humans do (indifferent vs. interactive). Gravity is the same regardless how often the stone falls, love not (constant vs. evolving). A stone need not learn how to fall, humans need to construct their personality (natural vs. constructing).

Now, it has happened—I used the word "construct". However, "real" and "construction" are not meant as antipodes here. As a statement about the nature of latent variables "real" and "constructed" exclude each other, turning to the social world they are entangled.

There are things of different kinds in our world (Hacking 1999). Natural kinds are in the world without any human involvement. Human kinds are created by humans; this is true of everything in the social world, be it norms, roles, identities, whatever. Human kinds can be created and are newly created. Hacking illustrates this point with considering Sartre's famous narration of how a French garçon de café attempts to fulfill the role of being a garçon de café.[4] This role took shape in the late 19th century; before it was not existent. This means that being such a garçon de café was not on option for a person in France just a few hundred years ago; it was simply not on the map of possibilities.

[4] Sartre describes the as follows: "His movement is quick and forward, a little too precise, a little too rapid. He comes toward the patrons with a step a little too quick. He bends forward a little too eagerly, his voice, his eyes express an interest a little too solicitous for the order of the costumer" (Sartre 1956, 59; quoted in, Hacking, 2002, 108-109).

Hacking says that by making up human kinds, we also make up people. The person working the tavern could not be garçon de café without the human kind "garçon de café". There are, of course more serious examples of making up people, e.g. mental illness (see Hacking 1999, 100ff.). The crucial point about human kinds is twofold. The available human kinds constitute my options to act and to identify. Once human kinds are used to classify people—as social science does—this changes not only the reality of the classified but the classified people will also change the meaning of the categorization. Hacking (2002) has called this the looping-effect of human kinds. This may sound strangely idealistic, as if coining a term would change reality. As Hacking himself says, a unique story probably needs to be told

about each human kind. Not all constructions are as contingent as the role of a garçon de café. Socal values, for example, are a necessary construction (Schwartz 1992).

In the same vein, Giddens (1984, 1987) sees social science as an endeavor of "double hermeneutics", "a mutual interpretative interplay between social science and those whose activities compose its subject matter" (Giddens 1984, xxxii). In the social sciences there is no way to keep the conceptual apparatus of the scientist away from the observed social reality; they circulate in and out of the social world they are coined to analyze.

Unlike a stone, social reality is socially constructed. Does it mean that social reality is less real? No, but it follows—and is evident—that social reality is dynamic. A realism adequate to the social word, therefore is a constructive realism (Dux 2000). On a methodological level, the question of the realness of latent variables prevails, albeit with a qualification: They are not intended to grasp a distant natural reality, but to re-construct the worldly constructions.

Social science as re-construction

A prerequisite of re-construction is to observe social reality. To find a data proxy for a social construction, we first need to understand the social construction itself. Key parameters in this undertaking are:

(1) contextuality

(2) disclosure

(3) necessity vs. interest

The first criterion is descriptive, the second is reasoning from description. The third criterion is most difficult as it depends on knowledge. *Contextuality* means that social constructions can change over time and vary across contexts. The more rapidly a concept changes over time, the more sensitive we need to be to keep track of the change in the way we set up our data proxies. The more we observe that a concept is bound to a specific social context, the more context bound our data proxies must be. *Disclosure* is an implication of contextuality. It highlights that social constructions may or may not be salient across situations. To put it in concrete terms, measurement models assume the latent variable to be the cause of the answer to the item. To work out, the latent variable needs to be salient during the response process. Answering a questionnaire is a specific situation. If the latent variable qua social construction is bound to another situation, it may not be salient in the interview situation. The goal is to trigger the latent variable in the situation of data collection. Questionnaires try to bring the latent variable into the situation by a triggering item formulation. For some social constructions, however, surveys might not be capable of achieving disclosure and other modes of data collections might be in order.

Necessity vs. interest tries to highlight that there is continuum of social constructions from absolutely necessary to purely interest based. This is a tricky issue—though an issue social scientist cannot avoid. Traditionally the issue has been framed as the value problem or the postulate of value freedom. The discussion stays controversial. But it has became clear, that striving for objectivity does not lead to a value free social science (Becker 1967, Gouldner 1963); be it for the very reason that the social sciences are a "subject-subject-relation" (Giddens 1984: xxxii). Taking advantage of the increase in knowledge in historical sociology, anthropology, and biology (for an intergrated theory, see Dux 2000), I propose a differentiated view. The reconstruction of phylogenesis (i.e. the development of mankind) and of ontogenesis (i.e. the development of every individual) allows to figure out which social constructions seem to develop out of necessity and which social constructions seem to serve particular interests. Developing a basic sense of morale, for example, is an anthropological necessity for every human, in order to survive in a social world he is completely dependent on. Developing a concept of the nation as a real community of faith is not. The living proof is Benedict Anderson who holds that nations are imagined communities (Anderson 1983). Obviously, Anderson did not follow the social construction of the nation as an ethnic group but decided to de-construct it. On

which grounds could he do that? I have state that all in social reality is constructed—so why applaud to the de-construction of a specific aspect? Anderson exposed that the prevailing construction of the nation as an ethnic group employed a naturalism that is not true. Nationalism treats ethnicity as a natural kind; Anderson as the social kind which it is. Facing interest-based claims that show up as quasi-natural kinds, the social scientist can and should take a position. This is science as enlightment.

A social scientist should also know that it is not done with de-construction. Showing that community is imagined does not make it imaginative (Jenkins 1996, 2008); it only puts nations to its proper place as social constructions. Yet, this is valuable, in that it opens discussion. It is valuable because the social scientist knows that there can be and that there are different conceptions of the nation. The debate over the nation is, of course, normative. Nothing speaks against engaging in a normative discourse, as long as the scientist abides to his own scientific findings, i.e. does not himself an unwarranted naturalism. I have tried to add my two cents to de-construct the idea of a European demos proposed a non-naturalist foundation of European integration as social integration (Datler 2012).

In many cases it will be debatable if a social construction is interest based or anthropologically necessary. As I said necessity vs. interest is a continuum; but it does provide a suggestion which evidence to look for when scientist have to decide on which side they are on.

5. Two case studies: social values and European identity

Social values and European identity are two rather different human kinds. These two cases need different approaches to validation. In describing both social values and European identity in terms of contextuality, disclosure, and necessity vs. interest; I hope to make a point that can be useful for other human kinds.

CASE I: social values

In developing the case of social values I will largely follow the theory of basic human values (Schwartz 1992, 1994). When I take issue with his position, this will be on logical or on methodological grounds. In general, I will not pretend to be what he is, an expert in value research. I will be concerned with the value priorities of individuals, not with Schwartz's theory of the cultural value system (see, Schwartz 2006).

Schwartz calls his theory “a theory of basic human values”. It is a strong theory with precise assumptions: There is definite structure of values. This structure is universal, i.e. to be found around to world. He is able to arrive at this strong theory in an openly functionalist argumentation that draws on the structural-functionalist theory of Parsons (1991 [orig. 1951]). Functionalism holds that what you want to postulate must fulfill certain functions. The critical issue is, “for which purpose?” (even more critical is the question “Functional for whom?”). According to Schwartz, value priorities are responses to three universal requirements of human existence: biological needs, requisites for coordinated social interaction, and demands of group survival and functioning. As these requirements are universal, the underlying values are universal. Funtionalism has not an easy stand in correct sociological theory (for a critique see, e.g. Giddens), one of the main objections is that a functional argument is not definite and always favors the status quo. The very same function could also be fulfilled by another social arrangement. Schwartz seems relatively well equipped to counter such a critique for two reasons. First, he allows opposite value orientations that fulfil the same function. Second, he elaborated his theory in studies in different regions of the world. I would only like to tentatively raise one issue: Schwartz always arranges his values in circle, i.e. in two dimensional space (see figure 1, Paper 1). To establish and validate this structure, he used multi-dimensional scaling plots. These plots are two-dimensional per definition. I mention this issue because both studies we conducted found a correlational structure of the values that is not circular. There are several studies that take issue with the universal circumplex structure of the values (Steinmetz et al. 2012, Perrinjaquet et al . 2007).

A problem-though maybe an unavoidable one—is that at the moment both theory and measurement models are revised (Schwartz et al. 2012). While it is welcome to see theory and measurement theory evolve, it seems that the arguments sometimes result in stalement between theory and empirical findings—for instance, is something wrong with the theory of circumplexity or are the measures not able to grasp it? An illustrative case is the search for the number of distinct values. In the first, widely used version of the theory there were ten values. Another boost of value studies came with the inclusion of Schwartz value items in the ESS (European Social Survey); but the ESS scale was a shortened version of the original scale. Using factor models (instead of multidimensional scaling as Schwartz 1992) researchers could only find five to eight values (Davidov et al. 2008). In the latest version there are 19 values (Schwartz et al 2012, Cieciuch et al. 2014). I am not

sure whether the number of values is the most important issue and whether statistical finger exercises as third(!)-order-factor-models are helpful. If we take circumplexity serious; it seems that any number of values is ok—it is just as slicing a cake into pieces. A very real issue, on the contrary, is the assumption of circumplexity as such. Opposing values should be negatively related; the available measures, however, do not behave this way. There are strong correlations between similar values. But there are no negative correlations between opposing values.

This criticism should not be misunderstood. It is easy to criticize the Schwartz theory, because it is theory in the best sense, it is well specified and therefore falsifiable. In Paper 1we concentrated on the theoretical claim of the universality of the values. Especially when a theory proposes universal concepts, comparisons across contexts are of interest. Valid comparisons requires equal measures, otherwise differences could either be substantive or methodological artifacts. This idea has become known as the requirement of measurement invariance or measurement equivalence (Van de Vijver and Leung 1997, Davidov et al. 2014). Invariance means that the same latent variable must relate to the same set of observations in the same way in each group. Statistically, this means that the mathematical function that relates latent variables to the observations must be the same in each of the groups involved in the comparison (Meredith 1993). The issue we addressed is this: The mathematical function of a factor model assumes that the observed variables are continuous, while the items are ordered categorical. This could be an additional source of model misfit and blur the judgment of invariance. Drawing on pioneering work on categorical data and factorial invariance (Millsap and Yun-Tein 2004), we reassessed the invariance of the value measures across the ESS samples of Luxembourgh, Belgium, and The Netherlands; which Davidov and Schmidt (2007) held already done with a factor model that assumes continuous indicators. The results are largely in line across the two studies, suggesting that the ordered categorical nature of the data did not impact the invariance judgment. This result does of courses not generalize to other likert-type scales; the study should rather be read as an example of how to implement an invariance study when ordinality is an issue. However, it should not be concealed that for ordinal factor models, the behavior of fit indices is less well researched and rules of thumb for invariance assessment are not available. The most promising approaches, however, are not about comparing fit indices; they suggest judging the effect of violations of invariance on the parameters of interest (Oberski 2014).

Paper 2 compares the Schwartz theory with one of its most prominent rivals, the Inglehart theory. Again, the comparison is restricted to values at the individual level. At the individual level general values orientations promise to be causal explanations of more concrete attitudes and behavior. Our approach is in line with the framework of construct validity. In a first step, internal validity, i.e. the quality of the measurement instruments, is assessed. In a second step, external validity, i.e. the relationships of values with supposed antecedents and supposed consequences, is assessed. It is apparent that the Schwartz value theory explicitly pays attention to the measurement of values conceptualized as underlying causes, i.e. a realist stance on latent variables. Values are regarded as the causes of observed answers to evaluative statements. Inglehart, though he employed the common factor model, seems to construct values as summary dimensions of very different ingredients. Among the constituencies of his value variables are attitudes, beliefs, and behavior—some of which are regarded as measures of other concepts such as national identity. Therefore, in terms of internal validity understood as *measurement* Schwartz fares better.

The more tricky part of the comparison concerns external validity. The theory of construct validity suggests assessing the validity of a latent variable by deriving hypotheses on relations with other variables from the theory in which the latent variable is embedded. Though it proofed difficult in some cases, we came up with a system of hypotheses for the same set of external variables for both value conceptions. The outcome of the empirical investigation, however, is inconclusive—some hits for Schwartz, some hits for Inglehart, some errors for Schwartz, some errors for Inglehart. “Maybe it is too early to discard on theory or the other”, this is how we framed the result in a preliminary version of the paper. Though this sentence did not make into the final version, the paper still has the remark that both theories are in a developmental stage. If I dare make a prognosis: The theories will develop but the situation will not change considerably. It is a firm wish to get unequivocal results when comparing the predictive power of two theories; in the social sciences it remains a phantasm to apply Occam’s Razor und to decide between competing theories. There are too many unknowns in the equation: the measures and measurement models of the value variables, the measures and measurement models of the external variables, the judgment of statistical significance and substantive importance, to name but a few. Every step is prone to error and open for argument. An example: Inglehart values did better than Schwartz values in the explanation of political

activism. But the item “signing a petition” is part of the Inglehart values measure. In a similar fashion, one could take issue with every single result. A comparison of Schwartz and Inglehart theories that has something to say about validitiy should be more concerned with theoretical and methodological issues than with regression coefficients. Maybe it is too early to compute regression coefficients for validation purposes; maybe it is the wrong way all together. One of the issues is, whether values are conceived of and measured as underlying causes or as heuristic summaries.

CASE II: European identity

The case of European identity is different compared to the case of values. For values, theories with an affinity to latent variables and issues of validity are available. For European identity an abundance of theories are available; but the mass of them pays little attention to variables and measurement. Often it is not their intention to treat European identity as a variable. Parallel to this strong philosophical and normative scholarship, empirical studies such as the Eurobarometer tried to produce numbers that are useful. Among survey researcher there evolved a consensus to take questions on the identification with Europe as measures of the European identity. However, this approach never was without its critics (Díez Medrano 2010, Díez Medrano and Gutiérrez 2001).

In the literature there are is only a single study that models European identity as a latent variable (Fuß 2006). One reason may be the lack of data sources with multiple indicators, but I suspect another reason. Buried in the files, I have a factor analytic study on the invariance of four measures of identification with Europe. The results in terms of invariance are meager. I could imagine that there are more of these studies in other drawers. More computing did not bring me closer to know why; so I stopped. I had to begin with theory.

Paper 3 starts out with the claim that the concept of European identity is overused but underspecified. European identity is not only used a lot but that European identity is used for very different purposes. European identity is an example par excellence of the looping effect of human kinds (Hacking 1996). In the political and philosophical discourse it is used to put forward normative claims, i.e. claims on what Europe should be. It can stand for post-national constitutional patriotism (e.g., Habermas 1998) or cosmopolitan orientation (e.g., Beck and Grande 2004); but is also used to promote Christianity (Frattini 2010). At the same time, same social scientists try to “measure” European

identity with questions on the degree of identification with Europe. Measure of identification with Europe, however, do not tap the contents of European identity. Exactly this lack of contents makes the resulting numbers susceptible to any normative (ab)use. This is not social science as reconstruction, this is social science as the fig leaf of identity politics.

In terms of the *contextuality*, European identity is very context dependant concept. In terms of *disclosure,* it is rather doubtful, if all aspects of European identity can be triggered in an interview. Identity is intrinsically social and more likely to be triggered in so-called inter-group situations (Jenkins 1996). European identity is clearly an *interest*-based social construction, there is no necessity for a European identity, but it could be useful for the European integration project. A useful conception of European identity, we concluded, must take into account the content of an identity Abdelal et al. (2009).

Paper 4 both works out the suggestion to take into account European identity content more systematically and puts it into practice. The approach acknowledges that we can measure the degree of identification with Europe, but to grasp why someone identifies with Europe we need to know what Europe means to her. Therefore, identity content is the key feature when we want to understand European identity and its relation to other social phenomena. In a nutshell, this is the theoretical conception. Defining European identity by its meaning has severe methodological implications. We need to empirically grasp the meaning. The meaning is socially constructed. Therefore, a valid account of European identity is an exercise in re-construction. We re-construct something that has been constructed in the first place. The constructions of Europe we find in social reality are the referents of our latent variable proxies. They are constructed; but real. They are real, but constructed. The realness of the constructions implies that social sciences can identify them. Thier constructedness implies that they are not fixed but changeable and contextual. I hesitate to use the term “measurement” in this context, because measurement suggest that there is something stable and enduring. Measurement also suggests that we could find a measurement instrument that would work for all time. I prefer the term re-construction to highlight that social science is always a bit behind and has to catch up with the social reality of European identity. This means to that we should not strive for equivalent factor models that measure European identity across time and social contexts. I propose to combine two parts: (1) an indicator of identification with Europe and (2) a latent class model of identity content. The identification part can be seen as measurement

and as the component of the method that is fixed, i.e. used again and again in the same way. Here multiple indicators would make sense to increase precision. The content part is typological and needs to be re-adjusted again and again, depending on the actual constructions of European identity in social reality.

Conclusions: Social Science as Re-construction and Quantitative Methods

The approach of construct validity needs highly developed theories to work out. Indeed for the studies on the Schwartz theory of basic human values, it was possible to use strategies of construct validation, which work fine as long as theory and data modeling agree. But at certain points—the excitement over the number of values qua factors, for instance—the general issue of "construct" glimpse through: An equating of theoretical concepts and empirical proxies.

Comparing the theories of Schwartz and Inglehart along the lines of construct validity showed the limitations of the approach. The theories in the social sciences have not conformed to a nomological network in the past and they most likely will never do so. At least it is tedious to wait for better theories to come and to work on even more fancy statistical models in the meantime.

Latent variable models as measurement models need two things: (1) highly specific theories (2) a world of latent variables with causal force. If, and only if, these two conditions are fulfilled, they are a tenable approach to measurement. The discussion of social values and European identity has shown that there are different human kinds. Not all aspects of social reality can be re-constructed as latent variable measurement models. It has to be elaborated rather than assumed that a latent variable measurement fits social reality. If it does not fit, this is neither the end of the world nor necessarily the end of quantitative social science. It is only the end of the myths of measurement.

Epilogue

Kenneth Bollen, one of the great masters of latent variable modeling, once said:

> *It is impossible to date the first use of latent variables. The idea that observable phenomena are influenced by underlying and unobserved causes is at least as old as religion, where unseen forces affect real world events.*
>
> (Bollen 2002, 606)

This passage is unfortunate, because from Bollen's work it is apparent that he is into rather technical stuff and avoids metaphysics. But some latent variable followers indeed behave as if latent variable modeling were the Holy Grail. But not all regard magic as desirable. Some even see latent variable modeling as a hex:

> *Thurstone stole fire from the gods. (As a punishment they chained him to factor analysis.)*
>
> (Lumsden 1980, 7)

In any case, there are strong emotions at play. I would suggest that we first use the fire and burn the myths about construct and then give back the fire to the gods. Afterwards nothing speaks against using latent variable modeling again, though it will be a less magical experience.

References

Abdelal, R., Herrera, Y. M., Johnston, A. I. & McDermott, R. (eds.) (2009) *Measuring identity. A guide for social scientists.* Cambridge, New York: Cambridge University Press.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999) *Standards for Educational and Psychological Testing. 2nd ed.* Washington DC: American Psychological Association.

Bartholomew, D. J. (1987) *Latent variable models and factor analysis*. London: Griffin.

Beck, U. & Grande, E. (2004) *Das kosmopolitische Europa. Gesellschaft und Politik in der zweiten Moderne.* Frankfurt am Main: Suhrkamp.

Becker, H. S. (1967) Whose side are we on? *Social Problems* 14:239–247.

Block, N. J. (1976) Fictionalism, Functionalism and Factor Analysis. In: Buck, R. C., , Schaffner, K. F. & Suppe, F (eds.) *Synthese library, vol. 101, Proceedings of the biennial meeting of the Philosophy of Science Association*. Dordrecht: Reidel.

Bollen, K. A. (1989) *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A. (2002) Latent variables in psychology and the social sciences. *Annual review of psychology* 53:605–634.

Bollen, K. A. & Bauldry, S. (2011) Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods* 16:265–284.

Bollen, K. & Lennox, R. (1991) Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin* 110:305–314.

Borsboom, D. (2006) The attack of the psychometricians. *Psychometrika* 71:425–440.

Borsboom, D.,Cramer, A. O. J. , Kievit, R. A., Zand Scholten, A., & Franić, S. (2009) The End of Construct Validity., In: , Lissitz, R. W (ed.) *The concept of validity. Revisions, new directions, and applications.* Charlotte, NC: Information Age Publishing, pp. 135–170.

Borsboom, D., Mellenbergh, G. J. 6 van Heerden, J. (2003) The theoretical status of latent variables. *Psychological Review* 110:203–219.

Borsboom, D. (2005) *Measuring the mind. Conceptual issues in modern psychometrics.* Cambridge : Cambridge University Press.

Bridgman, P. W. (1927) *The logic of modern physics.* New York: Macmillan.

Carnap, R. (1937) Testability and meaning II. *Philosophy of Science* 4:1–40.

Carnap, R. (1936) Testability and meaning I. *Philosophy of Science* 3:419–471.

Cieciuch, J., Davidov E., Vecchione, M., and Schwartz, Sh. H. (2014) A Hierarchical Structure of Basic Human Values in a Third-Order Confirmatory Factor Analysis. *Swiss Journal of Psychology* 73:177–182.

Cliff, N. (1983) Some Cautions Concerning The Application Of Causal Modeling Methods. *Multivariate Behavioral Research* 18:115–126.

Crano, W. D. & Brewer, M. B (2002) *Principles and methods of social research*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Cronbach, L. J. & Meehl, P. E. (1955) Construct validity in psychological tests. *Psychological Bulletin* 52: 281–302.

Davidov, E. & Schmidt, P. (2007) Are values in the Benelux comparable? In: Loosveldt, G., Swyngedouw. M. & Cambré, B. (eds.) *Measuring meaningful data in social research*. Leuven: Acco, 373-386.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014) Measurement Equivalence in Cross-National Research. *Annual Review of Sociology* 40:55–75.

Davidov, E., Schmidt, P., & Schwartz, Sh. H. (2008) Bringing Values Back In: The Adequacy of the European Social Survey to Measure Values in 20 Countries. *Public Opinion Quarterly* 72:420–445.

Devitt, M. (1997) *Realism and truth*. Princeton: Princeton University Press.

Díez Medrano, J. & Gutiérrez, P. (2001) Nested identities: national and European identity in Spain. *Ethnic and Racial Studies* 24:753–778.

Díez Medrano, J. (2010) Unpacking European Identity. *Politique Européenne* 30:45–66.

Duncan, O. D. (1984) *Notes on social measurement. Historical and critical.* New York: Russell Sage Foundation.

Franco, F. (2010) Speech to the Council of Europe. retrieved from: http://www.coe.int/t/dc/files/pa_session/jan_2010/20100126_disc_frattini_en.asp [29.06.2011]

Fuß, D. (2006) *Europa als Quelle sozialer Identität*. Dissertation, University of Gießen, Germany.

Garrett, H. E. (1937) *Statistics in psychology and education*. New York: Longmans, Green.

Giddens, A. (1984) *The constitution of society. Outline of the theory of structuration*. Cambridge [Cambridgeshire]: Polity Press.

Gould, S. J. (1996*) The mismeasure of man*. New York: Norton.

Gouldner, A. W. (1962) Anti-Minotaur: The Myth of a Value-Free Sociology. *Social Problems* 9:199-213.

Guttman, L. (1971) Measurement as structural theory. *Psychometrika* 36:329–347.

Habermas, J. (1998) *Die postnationale Konstellation. Politische Essays*. Frankfurt am Main: Suhrkamp.

Hacking, I. (2004) *Historical ontol*ogy. Cambridge, Mass.: Harvard University Press.

Haig, B. D. (2013) The Philosophy of Quantitative Methods. In: Little, T. D. (ed.) *The Oxford handbook of quantitative methods*. Oxford [u.a.]: Oxford Univ. Press, 7–31.

Hempel, C. G. (1952) *Fundamentals of concept formation in empirical science.* Chicago: University of Chicago Press.

Hood, B.S. (2008) *Latent Variable Realism in Psychometrics.* Doctoral Dissertation, Indiana University, United States of America.

Jenkins, R. (1996) *Social identity*. London: Routledge.

Kane, M. T. (2006) Validation. In: Brennan, R. L. (ed) *Educational Measurement (4th Edition)*. Washington, DC: American Council on Education, 17–64.

Lee, N., John W. C., & Chamberlain, L. (2013) The MIMIC model and formative variables: problems and solutions. *AMS Review* 3:3–17.

Lee, S., & Hershberger, S. L. (1990) A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research* 25:313–334.

Lumsden, J. (1980) Variations on a theme by Thurstone. *Applied Psychological Measurement* 4:1-7.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993) The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin* 114:185–199.

Maraun, M. D. (2003) *Myths and confusions: Psychometrics and the latent variable model*. Unpublished manuscript, retrieved from: http://www.sfu.ca/~maraun/Mikes%20page%20Myths%20and%20Confusions.html [04.07.2014]

Maraun, M.D. & Peters, J. (2005) What Does It Mean That an Issue Is Conceptual in Nature? *Journal of Personality Assessment* 85:128–133.

Maraun, M. D.& Gabriel, S. M. (2013) Illegitimate concept equating in the partial fusion of construct validation theory and latent variable modeling. *New Ideas in Psychology* 31:32–42.

Margolis E. & Lawrence, S. (2011) Concepts. *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab at Stanford University, retrieved from: http://plato.stanford.edu/entries/concepts/ [022.08.2014]

Meredith, W. (1993) Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58:525–543.

Messick, S. (1989) Validity. In: Linn, R. (ed.) *Educational measurement (3rd edition)* Washington, DC: American Council on Education, 13–100.

Michell, J. (2000) Normal Science, Pathological Science and Psychometrics. *Theory & Psychology* 10:639–667.

Michell, J. (2009) Invalidity in Validity. In: Lissitz, R. W (ed.) *The concept of validity. Revisions, new directions, and applications.* Charlotte, NC: Information Age Publishing, 111–133.

Michell, J. (2008) Is Psychometrics Pathological Science? *Measurement: Interdisciplinary Research & Perspective* 6:7–24.

Millsap, R. E. & Yun-Tein, J. (2004) Assessing Factorial Invariance in Ordered-Categorical Measures. *Multivariate Behavioral Research* 39:479–515.

Newton, P. E. (2012) Clarifying the Consensus Definition of Validity. *Measurement: Interdisciplinary Research & Perspective* 10:1–29.

Oberski, D. L. (2014) Evaluating Sensitivity of Parameters of Interest to Measurement Invariance in Latent Variable Models. *Political Analysis* 22:45–60.

Parsons, T. (1991) *The social system*. London: Routledge.

Perrinjaquet, A., Furrer, O., Usunier, JC.,Cestre, G. & Valette-Florence, P. (2007) A test of the quasi-circumplex structure of human values. *Journal of Research in Personality* 41:820–840.

Popper, K. R. (1959) *The logic of scientific discovery*. London: Hutchinson & Sons.

Rigdon, E. E. (2012) Rethinking Partial Least Squares Path Modeling: In Praise of Simple Methods. *Long Range Planning* 45:341–358.

Rigdon, E. E. (2013) Lee, Cadogan, and Chamberlain: an excellent point . . . But what about that iceberg? *AMS Review* 3:24–29.

Schnell, R., Hill, P. B., and Esser, E. (2008) *Methoden der empirischen Sozialforschung*. München: Oldenbourg.

Schwartz, Sh. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lonnqvist, JE., Demirutku, K., Dirilen-Gumus, O., & Konty, M. (2012) Refining the theory of basic individual values. *Journal of personality and social psychology* 103:663–688.

Schwartz, S. H. (1992) Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology* 25, 1-65.

Schwartz, S. H. (2006) A Theory of Cultural Value Orientations: Explication and Applications *Comparative Sociology* 5:137–182.

Schwartz, S. H. (1994) Are there Universal Aspects in the Content and Structure of Values? *Journal of Social Issues* 50:19–45.

Sireci, S. G. (2009) Packing and unpacking sources of validity evidence: history repeats itself again. : Lissitz, R. W (ed.) *The concept of validity. Revisions, new directions, and applications.* Charlotte, NC: Information Age Publishing, 19–37.

Skrondal, A. & Rabe-Hesketh, S. (2004) *Generalized latent variable modeling. Multilevel, longitudinal, and structural equation models*. Boca Raton: Chapman & Hall/CRC.

Steinmetz, H., Isidor, R. & Bäuerle, N. (2012) Testing the Circular Structure of Human Values: A Meta-Analytical Structural Equation Modelling Approach. *Survey Research Methods* 6: 61-75.

Stevens, S. S. (1946) On the theory of scales of measurement. *Science* 103: 677–680.

van de Vijver, F. J. R. and Leung, K. (1997) *Methods and data analysis for cross-cultural research.* Thousand Oaks: Sage Publications.

Zumbo, B. D. (2009) Validity as contextualized and pragmatic explanation, and its implications for validation practice. In: Lissitz, R. W. (ed.),The concept of validity: Revisions, new directions, and applications. Charlotte, NC: Information Age Publishing, 65–82.

Paper 1

Testing the Invariance of Values in the Benelux Countries with the European Social Survey: Accounting for Ordinality

---

Co-authored with: Eldad Davidov, University of Zurich
Peter Schmidt, University of Marburg
Shalom H. Schwartz, Hebrew University of Jerusalem

Published as: Davidov, Eldad, Datler, Georg, Schmidt, Peter & Schwarz, Shalom H. (2011) Testing the Invariance of Values in the Benelux Countries with the European Social Survey: Accounting for Ordinality In: Davidov, Eldad, Schmidt, Peter and Billiet, Jacques (eds) Cross-Cultural Analysis: Methods and Applications. New York: Routledge, 149-171.

Testing the Invariance of Values in the Benelux Countries with the European Social Survey: Accounting for Ordinality

Abstract

In this chapter we test the comparability of the measurement of human values (Schwartz, 1992) in the second round (2004-5) of the European Social Survey (ESS) across three countries, Belgium, Netherlands, and Luxembourg, while accounting for the fact that the data are ordinal (categorical but ordered, i.e., nonmetric). A previous study (Davidov and Schmidt, 2007) established metric invariance for seven values across these three countries but scalar invariance only for three values: self-direction, stimulation, and the unified values universalism-benevolence. However, the previous study applied multiple-group confirmatory factor analysis (MGCFA) and assumed the use of normally distributed, continuous data. Values are measured in the ESS on a 6-point scale, and responses are not normally distributed. In this chapter we address the criticism of Lubke and Muthén (2004) that MGCFA is not appropriate for testing invariance of Likert-type scales. We use a model for ordinal (ordered-categorical) indicators to test for measurement invariance. The general conclusions are consistent across these two methods.

1. Introduction

The increasing importance of comparative studies across countries and over time has encouraged the collection of survey data in diverse contexts and time points in recent decades (e.g., the European Social Survey, the International Social Survey Program, the European Value Study, or the World Value Survey). These surveys share the goals of collecting comparable responses from large, national representative samples and of gathering data at multiple points in time to permit the study of differences and similarities among cultures and change over time. The methodological literature has emphasized, however, that comparisons between groups and/or time points are not legitimate without first assessing whether the concepts used (e.g., human values) are indeed comparable across countries or over time (e.g., Billiet, 2003; Cheung and Rensvold, 2002; De Beuckelaer, 2005; Steenkamp and Baumgartner, 1998; Vandenberg, 2002; Vandenberg and Lance, 2000).

Several techniques have been developed to assess the comparability of concepts. Two of the most common techniques are multiple-group confirmatory factor analysis (MGCFA: Bollen, 1989; Jöreskog, 1971) and means and covariance structure analysis (MACS: Sörbom, 1974, 1978). These techniques can test for configural, metric, and scalar invariance (see, e.g., the chapters by Lee et al., Oreg et al., Read et al., Billiet and Meuleman, and Steinmetz in this volume). Configural invariance indicates that the same indicators measure the same theoretical constructs across groups or time points. Metric invariance is more restrictive; it indicates that respondents interpret the intervals on the response scale in a similar way across groups. Metric invariance with continuous latent variables indicators means that the loadings of the indicators on the factors are equal across groups and/or time points. This implies that the constructs tap the same content across the groups. The most restrictive level of invariance with continuous latent variables indicators, scalar invariance, requires that the intercepts of each item be the same across groups and/or time points. This means that respondents in different contexts use the same scale origin. As other chapters in this book explain (e.g., Billiet and Meuleman; De Beuckelaer; Oreg et al.; Steinmetz), metric invariance permits the comparison of correlates across countries and/or time. Scalar invariance also permits the comparison of latent variable means. The scalar invariance model constrains the means of the latent variables to zero in one group (referred to as the reference group) and estimates them in the other groups.

A differentiation can also be made between full and partial invariance. A partial metric invariance model constrains the factor loadings of at least two indicators of a construct to be equal across groups. A partial scalar invariance model constrains the factor loadings *and* intercepts of at least two indicators to be equal across groups (Byrne et al., 1989; Steenkamp and Baumgartner, 1998). If the factor loadings of *all* indicators are constrained to be equal across groups, we term the model full metric invariance. If the intercepts *and* factor loadings of *all* indicators are constrained to be equal across groups, one terms the model full scalar invariance.

MGCFA is designed for continuous and normally distributed data. Nevertheless, it is often applied to Likert-type scales where researchers typically assume continuity and a normal distribution underlying the scales used. However, Lubke and Muthén (2004) have criticized the analysis of Likert-type scales under the assumption of multivariate normality. If Likert-type data is analyzed assuming that multivariate normality holds, different factor structures may be found in different groups even if these factor structures are actually invariant across

groups (however, they find that the estimates of latent mean differences are rather robust). They propose, instead, fitting a model for ordinal indicators.

The structural equation modeling software programs LISREL (Jöreskog and Sörbom, 1996) and Mplus (Muthén and Muthén 2007) use two different strategies to identify a model which is fitted to ordinal data. They both include threshold parameters and polychoric correlations between the measurements. The major difference between the two programs is that the LISREL program assumes that the thresholds are equal across groups whereas the Mplus program allows the actual testing of whether this is so. Based on this difference we have chosen to apply the Mplus approach to the data presented in this chapter.

Several simulation studies have shown that MGCFA works well when testing for cross-cultural invariance even when the data are ordinal rather than continuous or normally distributed (Welkenhuysen-Gybels and Billiet, 2002; Welkenhuysen-Gybels, 2004; De Beuckelaer's chapter in this volume).[1] The current chapter is, to the best of our knowledge, the first to compare the results of an invariance test using MGCFA for ordinal data with the outcomes of an MGCFA that assumes continuous indicators using actual survey data on basic human values available from the European Social Survey (ESS).

The following section first describes the theory of human values that we will assess. Then we provide a description of the method that is utilized for measuring values in the ESS. In the empirical part of the chapter, the results of measurement invariance testing using MGCFA under the assumption of normality reported by Davidov and Schmidt (2007) are summarized. This is followed by a presentation of the procedure for testing invariance designed for ordinal indicators and the results of applying this method to the same data. We conclude with summarizing remarks and considerations.

2. Theory

Values have played an increasingly important role in the social sciences in recent decades. However, the absence of a widely accepted theory for conceptualizing values and of valid scales to measure them have limited researchers' ability to conduct empirical studies using the

[1] These studies report simulations which examine whether assuming normality and continuity of measurement scales when using ordinal categorical scales yields different conclusions in a cross-cultural invariance test. Comparisons of several estimation methods based on different assumptions for other types of models have also been conducted. They generally conclude that the maximum likelihood parameter estimates and standard errors are rather robust for small violations of normality (see e.g., Coenders, Satorra and Saris, 1997; and Coenders and Saris, 1995).

value concept. This changed with the introduction of the Schwartz (1992) value theory. This theory specifies 10 basic values that form four higher-order value dimensions that people around the world apparently recognize (Schwartz, 1992, 1994, 2005). Starting with its first round in 2002-3, the ESS included an instrument to measure the 10 values in the theory.

The theory defines values as *desirable, trans-situational goals, varying in importance, that serve as guiding principles in people's lives.* It proposes 10 motivationally distinct human values: power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity, and security. Table 1 presents the motivational goal that each value expresses. For example, the core motivational goal of power values is social status and prestige, control or dominance over people and resources.

Table 1: Definitions of the motivational types of values in terms of their core goal

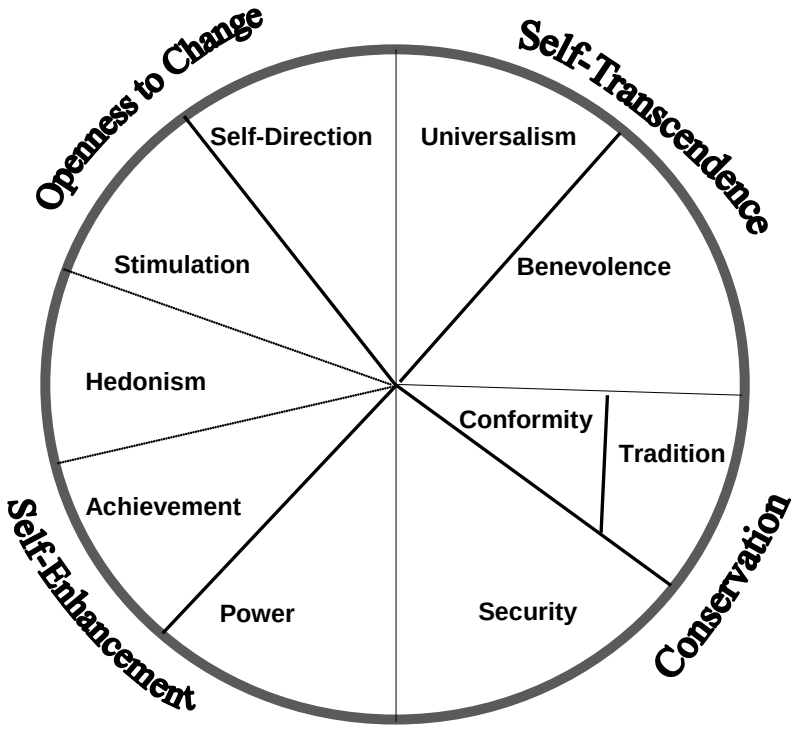
POWER	Social status and prestige, control or dominance over people and resources
ACHIEVEMENT	Personal success through demonstrating competence according to social standards
HEDONISM	Pleasure and sensuous gratification for oneself
STIMULATION	Excitement, novelty, and challenge in life
SELF-DIRECTION	Independent thought and action-choosing, creating, exploring
UNIVERSALISM	Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature
BENEVOLENCE	Preservation and enhancement of the welfare of people with whom one is in frequent personal contact
TRADITION	Respect, commitment, and acceptance of the customs and ideas that traditional culture or religion provide the self
CONFORMITY	:Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms
SECURITY	Safety, harmony, and stability of society, of relationships, and of self

* Adopted from Sagiv and Schwartz (1995).

The theory also postulates dynamic relations of interdependence among the values. Values are compatible if the motivational goals they express can be pursued simultaneously (e.g., conformity and tradition). Such values will correlate positively with each other. Values are incompatible if pursuing the motivational goal of one value conflicts with pursuing the goal of the other (e.g., security and stimulation). Such values will correlate negatively. Values are neither compatible nor incompatible if pursuing the motivational goal of one does not affect the other (e.g., benevolence and self-direction). Such values will typically show nonsignificant correlations.

Figure 1 portrays the full set of dynamic relations among the 10 values. Values with compatible motivational goals are close to each within the circle, those with incompatible goals are further apart. For example, the values power and universalism are far apart from each other within the circle. This reflects the theoretical idea that pursuing the goals of universalism values tends to oppose pursuing the goals of power values. Devoting oneself to the welfare of *all* people is largely incompatible with seeking control and dominance for oneself over other people and resources. In contrast, universalism and benevolence are adjacent to each other in Figure 1 because their goals are compatible: It is possible to devote oneself to the welfare of all and also to seek to enhance the welfare of those with whom one is close.

Figure 1: The dynamic relations between the values



The theory distinguishes 10 value factors, but one may not always be able to discriminate all 10 values empirically. This may occur, for example, if there are practical measurement restrictions such as very few items to measure each value and poor discriminant validity between values with related motivational goals (Campbell and Fiske, 1959; Knoppen and Saris, 2009). Instead, pairs of adjacent values (e.g., universalism and benevolence) may be captured as a single value (e.g., a unified universalism-benevolence value).

The compatibilities and oppositions among the values may be summarized using the higher order factors and dimensions shown in Figure 1. One dimension contrasts self-enhancement and self-transcendence values. This dimension contrasts power and achievement values (that emphasize one's pursuit of success and dominance) to universalism and benevolence values (that involve concern for the welfare and interests of other people). The second dimension contrasts openness to change and conservation values. It opposes self-direction and stimulation values (that emphasize independence and readiness for new experiences) to conservation values (that emphasize self-restriction, order, and resistance to change) (for a more detailed discussion see, e.g., Schwartz, 1992, 1994).

3. The ESS Measurement of the Ten Basic Human Values

A shortened version of the original 40-item portrait value questionnaire (PVQ) to measure values (Schwartz, Melech, Lehmann, Burgess, Harris, and Owens, 2001; Schwartz, 2005) was developed for the ESS. Due to time and budgetary constraints, the ESS instrument includes only 21 questions to measure the 10 values. Two items were chosen for each value (three for universalism) with the objective of providing maximum coverage of the conceptual breadth of the value rather than to maximize internal indicator reliability and high discriminant validity.

The ESS scale describes 21 different people, gender-matched with the respondent. Each description portrays a person in terms of what is important to him or her, thereby pointing to one of the 10 values. For example: "Thinking up new ideas and being creative is important to him. He likes to do things in his own original way" describes a person for whom self-direction is important. Regarding each description of a person, respondents are asked to answer: "How much like you is this person?" Responses are recorded on a Likert-type rating scale ranging from 1 (*not like me at all*) to 6 (*very much like me*). Respondents' own values are inferred from their self-reported similarity to people described implicitly in terms of particular values. Table 2 presents the 10 values as they appear in the ESS scale. Two items measure each value with the exception of universalism which is measured by three items because of its very broad content.

Table 2: The ESS Human Values Scale in the 2nd round (male version)

Value	Item # (Numbered and Labeled as in the ESS Questionnaire)
Self-direction (SD)	1. Thinking up new ideas and being creative is important to him. He likes to do things in his own original way (ipcrtiv). 11. It is important for him to make his own decisions about what he does. He likes to be free to plan and not depend on others (impfree).
Universalism (UN)	3. He thinks it is important that every person in the world be treated equally. He believes everyone should have equal opportunities in life (ipeqopt). 8. It is important for him to listen to people who are different from him. Even when he disagrees with them, he still wants to understand them (ipudrst). 19. He strongly believes that people should care for nature. Looking after the environment is important to him (impenv).
Benevolence (BE)	12. It is very important for him to help the people around him. He wants to care for their well-being (iphlppl). 18. It is important for him to be loyal to his friends. He wants to devote himself to people close to him (iplylfr).
Tradition (TR)	9. It is important for him to be humble and modest. He tries not to draw attention to himself (ipmodst). 20. Tradition is important to him. He tries to follow the customs handed down by his religion or his family (imptrad).
Conformity (CO)	7. He believes that people should do what they're told. He thinks people should follow rules at all times, even when no one is watching (ipfrule). 16. It is important for him to always behave properly. He wants to avoid doing anything people would say is wrong (ipbhprp).
Security (SEC)	5. It is important for him to live in secure surroundings. He avoids anything that might endanger his safety (impsafe). 14. It is important for him that the government insures his safety against all threats. He wants the state to be strong so it can defend its citizens (ipstrgv).
Power (PO)	2. It is important for him to be rich. He wants to have a lot of money and expensive things (imprich). 17. It is important for him to get respect from others. He wants people to do what he says (iprspot).
Achievement (AC)	4. It is important for him to show his abilities. He wants people to admire what he does (ipshabt). 13. Being very successful is important to him. He hopes people will recognize his achievements (ipsuces).
Hedonism (HE)	10. Having a good time is important to him. He likes to "spoil" himself (ipgdtim). 21. He seeks every chance he can to have fun. It is important for him to do things that give him pleasure (impfun).
Stimulation (ST)	6. He likes surprises and is always looking for new things to do. He thinks it is important to do lots of different things in life (impdiff). 15. He looks for adventures and likes to take risks. He wants to have an exciting life (ipadvnt).

Source: Adapted from Davidov, E. Survey Research Methods, 2(1), 33–46, 2008.

4. Empirical Analyses

Three countries are included in the analysis: Belgium (N = 1,778), Luxemburg (N = 1,635) and the Netherlands (N = 1,881) (total N = 5,294). Details on the data collection techniques that were used in each country are documented on the website http://www.europeansocialsurvey.org. The data used in the analyses were downloaded from the website http://ess.nsd.uib.no.

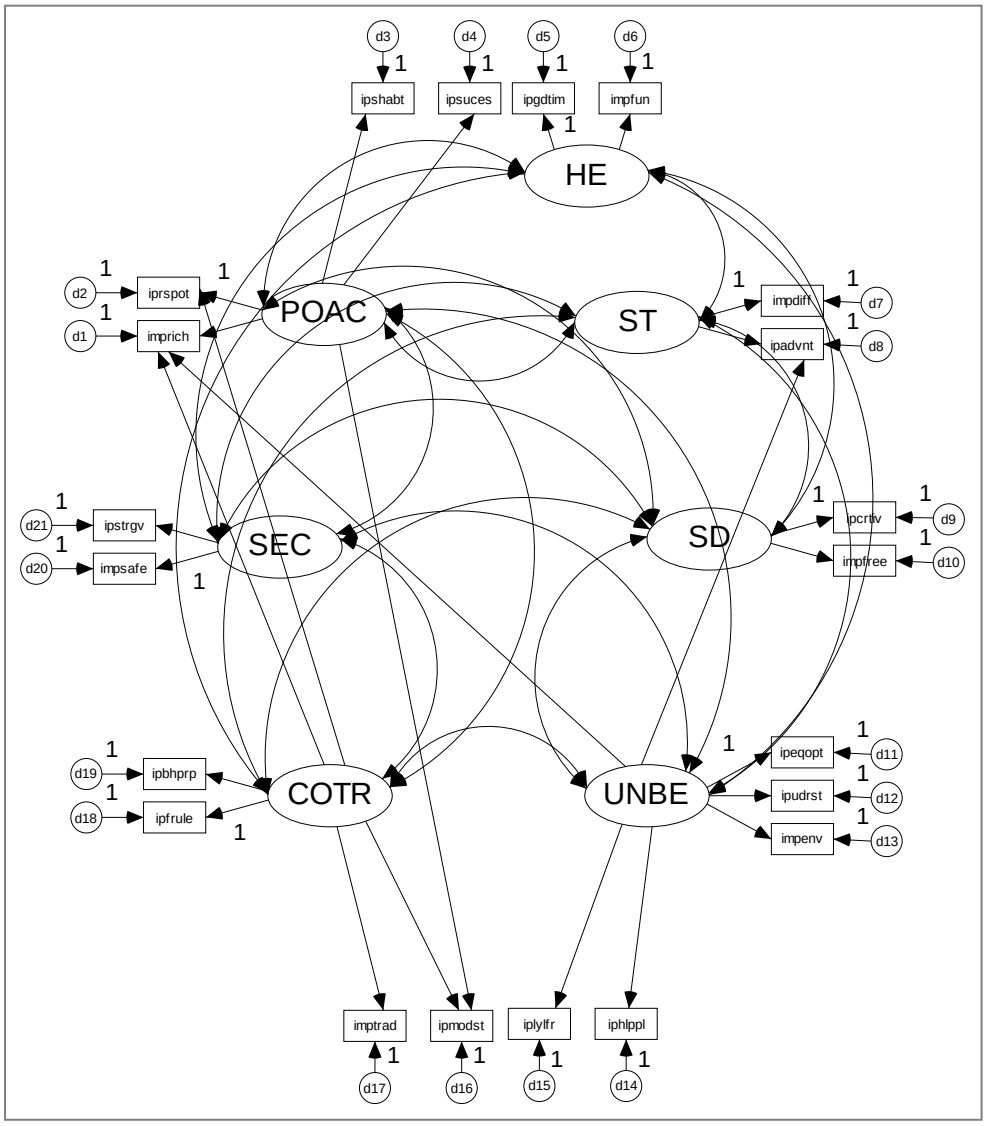
4.1 Previous Findings

The first studies to assess invariance of the measurement of values across ESS countries applied MGCFA (Davidov, 2008; Davidov and Schmidt, 2007; Davidov, Schmidt and Schwartz, 2008). In these studies, seven distinct values, rather than the 10 values postulated by the theory, were identified in most of the ESS countries. Three pairs from the original 10 values had to be unified: power with achievement, conformity with tradition, and universalism with benevolence. These pairs of values had very high intercorrelations and could not be modeled separately. Two reasons have been proposed for this finding. First, the use of 21 instead of the original 40 PVQ questions to measure values does not provide a sufficient number of questions to measure each value separately in a confirmatory factor analysis (CFA) (Davidov, Schmidt and Schwartz, 2008). Second, high correlations among some values attest to a lack of discriminant validity (Knoppen and Saris, 2009), requiring them to be unified.

The three pairs of values that had to be unified are adjacent in the circular theoretical structure portrayed in Figure 1. The need to unify them, therefore, does not contradict the theory of the circular value structure. In addition, five additional paths (cross loadings) were introduced in the MGCFA to improve the solution. Each addition was a path from one of the unified value factors to a distant value indicator: (1) from the universalism-benevolence factor to the item 'important to be rich', (2) from the universalism-benevolence factor to the item 'important to have adventures', (3) from the conformity-tradition factor to the item 'important to get respect from others', (4) from the power-achievement factor to the item 'important to be modest', and (5) from the conformity-tradition factor to the item 'important to be rich' (Davidov, Schmidt

and Schwartz, 2008). [2] Figure 2 depicts the best fitting model from this study that was calculated using the structural equation modeling software Amos (Arbuckle, 2005).

Figure 2: The CFA model



All values are allowed to correlate with each other. The large circles represent the values. For example, HE represents the value hedonism, and UNBE represents the unified value universalism-benevolence. The small circles represent measurement errors. The rectangles stand for the indicators measuring the values. For item and value abbreviations see Table 2.

[2] The negative cross loadings indicate that the association (covariance) between the opposing latent value constructs did not capture all of the opposition for these items. The positive cross loadings indicate that these associations overestimated the opposition for two items. The need for these cross loadings may be due to the reduction from 10 original values to seven. Without introducing them, the model fit was not acceptable. From a measurement point of view, cross loadings are not elegant. Cross loadings contaminate correlations between factors, a problem if one is interested in the correlations. However, our main interest was not to evaluate the strengths of relationships between values but to examine whether measurement properties, such as factor loadings and intercepts, are invariant across countries.

Davidov and Schmidt (2007) tested for invariance of the value measurements across three countries: Belgium, Luxembourg, and the Netherlands. They used the model just described (that differentiates between seven values) in their application of MGCFA to data from the second round of the ESS. In their analysis they assumed continuity and normality of the value scales. They reported that configural and metric invariance was found across the three countries for the seven values. However, analysis of the data did not support scalar invariance for all seven values. Rather, only stimulation values, self-direction values, and the unified universalism-benevolence value displayed scalar invariance. Thus, latent mean comparisons between the three countries are legitimate only for these three values (Steenkamp and Baumgartner, 1998).

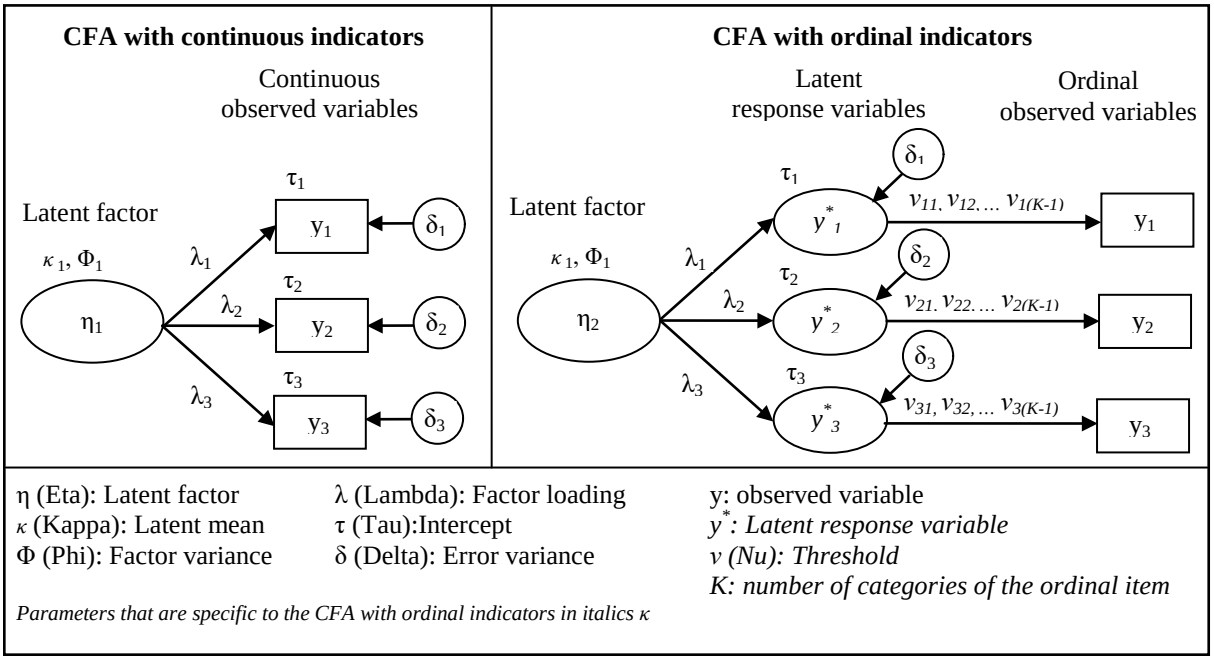
In the present analysis we assess the invariance of the value scales when fitting the model for ordinal indicators using MGCFA. To do this, we employ the same value data for Belgium, Luxembourg, and the Netherlands available from the second round of the ESS. Finally, we compare the results of this test with the findings of Davidov and Schmidt (2007).

4.2. Testing for Invariance While Accounting for Ordinality of the Data

Muthén, du Toit, and Spisic (1997) and Muthén and Muthén (1998, pp. 357–358) propose a theoretically more appropriate method for ordinal (ordered-categorical) scales. The proposed method fits a CFA model to polychoric correlations using robust weighted least squares (robust WLS) (see also Flora and Curran, 2004). This approach is based on the work of Satorra and colleagues (Chou, Bentler and Satorra, 1991; Satorra, 1992; Satorra and Bentler, 1990). This estimator is available in the software program Mplus. In the present analysis we use Mplus version 3.0 (Muthén and Muthén, 2007) to assess the invariance of the value measurements, following procedural guidelines suggested by Millsap and Yun-Tein (2004) and Temme (2006).

Figure 3 displays the path diagram of a single factor CFA for the ordinal case compared to the continuous case. In contrast to continuous indicators, CFA (and MGCFA) for ordinal data (such as Likert-type scales) assumes that the observed items (y's in Figure 3) are not directly influenced by their corresponding latent factor but indirectly via a continuous latent response variable (y^* in Figure 3) (Temme 2006). The main difference resulting from this specification is that we have to estimate item-specific threshold parameters (ν in Figure 3).

Figure 3: CFA with ordinal indicators compared to CFA with continuous indicators



These threshold parameters (ν in Figure 3) partition the continuous normally-distributed latent response variable into several categories. With six response categories for each item in our case there are five thresholds. [3] If the value for the continuous latent response variable exceeds a threshold, the observed value of the item changes to the next category. The ordinal CFA model, just like the continuous CFA model, contains factor loadings (λ in Figure 3) and intercepts (τ in Figure 3). Whereas in the continuous case factor loadings and intercepts are parameters of the observed indicators, in the ordinal case factor loadings and intercepts are parameters of the latent response variable (for further details, see Muthén and Asparouhov, 2002; Millsap and Yun-Tein, 2004).

In testing for invariance with continuous MGCFA, the distinction between metric invariance and scalar invariance is well established. In the continuous case it is *only the factor loadings* that determine the slopes of the item response curves. The intercepts only influence the starting points of the item response curves but not their slopes. Therefore, it is sufficient to constrain the factor loadings to be equal to guarantee metric invariance and to compare structural associations of the latent variables across groups. Latent mean comparison requires equality of intercepts (scalar invariance) in addition.

[3] A threshold captures transitions from one category to another. Thus, if there are K response categories for an indicator, there are K-1 threshold parameters for the latent response variable.

In the *ordinal* CFA (see Figure 3), in contrast, the item probability curves (i.e., the scores of the ordinal indicators) are jointly influenced by the factor loadings (λ's), the intercepts (τ's), and the thresholds (ν's). Thus, guaranteeing only that factor loadings are equal across groups is not enough to ensure that the item response curves are comparable in the ordinal case. Comparison of group means is still not permissible. Establishing measurement invariance in the ordinal case requires constraining factor loadings, thresholds, and intercepts simultaneously. Thus, a distinction between metric and scalar invariance is not substantively meaningful in the ordinal case because there is only one step in the measurement invariance test, the step which constrains all parameters to be equal.

Identifying the measurement invariance model for the ordinal case requires a somewhat different set of constraints than those required in the continuous case (Millsap and Yun-Tein, 2004; Temme, 2006). The reason is that it is not possible to estimate the thresholds and the intercept at the same time. The programs LISREL and MPLUS employ different strategies to deal with this issue.

The LISREL program (Jöreskog and Sörbom, 1996) constrains all the threshold parameters to be equal across groups. This allows testing the equality of the intercepts and factor loadings. However, there is no reason to believe that thresholds are equal. The equality of thresholds is an empirical question that can only be answered in an empirical test. The Mplus program (Muthén and Muthén, 2007) permits testing the equality of threshold parameters but constrains, as a default, all the intercepts to be zero (for identification purposes). Estimating the thresholds and constraining the intercepts is more informative than the other way around. In fact, the intercept parameter is equal to a constant shift in all the thresholds of an indicator.[4] If intercepts are found to be different across groups, one does not know whether this is due to differences in the full set of thresholds across groups or whether only some of the thresholds differ across groups. By contrast, testing for the equality of thresholds may allow certain thresholds to differ across groups. Thus, the flexibility of Mplus makes it preferable to LISREL for testing invariance in the ordinal case (Millsap and Yun-Tein, 2004).

The minimum number of constraints required to identify the model in the ordinal case depends on the number of categories of the observed variables and the model structure. Two model structures can be considered, a congeneric model and a noncongeneric model (Millsap

[4] In principle, the equality constraint of the intercepts may be released in Mplus by introducing a perfectly measured factor behind the latent response variable (Muthén and Asparouhov, 2002, p. 15).

and Yun-Tein, 2004). In a congeneric model there are no cross loadings, and each item loads only on one factor. In a noncongeneric model one or more items may load on more than one factor. Accordingly, in the congeneric model one has to set at least one threshold for each indicator to be equal across groups. In addition, a second threshold of the reference indicator of each latent variable is also constrained to be equal across groups. In the noncongeneric model, one has to constrain at least two thresholds to be equal for all indicators. In this model all intercepts have also to be constrained to zero. Our model includes five cross loadings and is thus noncongeneric. In addition, as in the continuous case, the factor loading of one reference indicator has to be set to 1.

Testing for invariance in the ordinal case implies setting *all* intercepts and thresholds equal across groups. However, it is not possible to test, at the same time, whether intercepts *and* threshold parameters are equal. One can either test for the equality of thresholds assuming that all intercepts are equal, or test for the equality of intercepts assuming that all thresholds are equal. Testing for the equality of thresholds is more informative than testing for the equality of intercepts. Thus, in line with the Mplus default (see Muthén and Asparouhov, 2002), we constrained intercepts to zero and all thresholds and factor loadings to be equal across groups.[5]

We adopted a top-down strategy. We started with the most restrictive model which imposes equality constraints across groups on the thresholds, intercepts, and factor loadings. Then we gradually released some of the equality constraints on the thresholds for the indicators whose constructs did not pass the scalar invariance test in the continuous model. To decide whether the data support a model, we followed the cut-off criteria suggested by Hu and Bentler (1999) and Marsh, Hau, and Wen (2004). They suggest a minimum value of 0.90-0.95 for the comparative fit index (CFI), a global fit index and a maximum value of 0.05-0.08 for the root mean square error of approximation (RMSEA) fit measure. To compare models, we used the criteria proposed by Chen (2007) who suggested assessing differences between models by looking at differences in the fit measures CFI and RMSEA. If both the decrease in the CFI and the increase in the RMSEA in a more restrictive model are smaller than 0.01, then the

[5] Two parameterizations are possible for running the model: Theta and Delta (Muthén and Muthén, 2007). The Theta parameterization includes residual variances for the continuous latent response variables (see Muthén and Muthén, 2007, pp. 485-486). This has the advantage of also permitting a test of the invariance of the residual variances (Millsap and Yun-Tein, 2004, Muthén and Asparouhov 2002). We applied both the Theta and Delta parameterizations and obtained essentially the same results. See the appendix for the final model.

more restrictive model can be considered as acceptable. We do not use the chi-square difference test and the p-value to distinguish between models because even small misspecifications may lead to model rejection with large sample sizes (Cheung and Rensvold, 2002). [6]

4.3 Results

Table 3 Global fit measures in the different models[a]

	MGCFA Under Assumption of Normality—the Continuous Case	MGCFA Using Robust WLS—the Ordinal Case
Model 1: Full measurement invariance (scalar in the continuous case) of seven values		
Chi-square	4165	3888[b]
df	555	355[b]
p-value	.000	.000[b]
RMSEA	.035	.077
PCLOSE	1.000	—[c]
CFI	.849	.855
Model 2: Full measurement invariance (scalar in the continuous case) of three values (UNBE, ST, and SD)		
Chi-square	2926	2492[b]
df	539	310[b]
p-value	.000	.000[b]
RMSEA	.029	.065
PCLOSE	1.000	—[c]
CFI	.900	.911

[a] df = degrees of freedom; RMSEA = root mean square error of approximation; PCLOSE = probability of close fit; CFI = comparative fit index; For details see, for example, Arbuckle, J. L., *Amos 6.0 User's* Guide, SPSS, Chicago, IL, 2005 and Muthén, L. K., and Muthén, B. O., *Mplus User's Guide*, Muthén & Muthén, Los Angeles, CA, 2007.

[b] With WLSMV estimation (which corresponds to robust WLS), chi-square values are mean and variance adjusted and degrees of freedom are estimated rather than derived from the model structure. Chi-square and the number of degrees of freedom are thus not meaningful themselves but rather adjusted in order to provide correct p-values. The formulas are available in the Mplus technical appendices at www.statmodel.com.

[c] PCLOSE is not provided by Mplus for multiple group analyses.

Table 3 summarizes the global fit measures of the models—chi-square, degrees of freedom (df), p value, RMSEA, p of close fit (PCLOSE), and CFI. Column 1 presents the fit measures using MGCFA under the assumption of multivariate normality (based on Davidov and Schmidt, 2007). We name this model 'the continuous case' model. Column 2 presents the fit

[6] Before testing for measurement invariance, we examined the level of skewness and kurtosis of the values across the countries. Skewness (both left and right, depending on the item) was significant for all 21 items in the three countries. Kurtosis was significant for 20 items.

measures after accounting for ordinality of the outcomes. We name this model 'the ordinal case' model.

Table 3 clearly shows that the most restrictive (invariant) model (Model 1), for seven values, is rejected by the data in both the continuous and ordinal case. Whereas the RMSEA displays satisfactory levels, the CFI does not achieve the minimum criterion for an acceptable fit. Model 2 constrains invariance for only the three values that Davidov and Schmidt (2007) reported as showing scalar invariance: stimulation (ST), self-direction (SD), and the unified value universalism-benevolence (UNBE). The measurement of the four other values hedonism (HE), security (SEC), and the unified values conformity-tradition (COTR), and power-achievement (POAC) is allowed to vary across countries. As shown, the data supported this model both in the continuous and in the ordinal case. In other words, it is legitimate to compare the means of these three values across countries.

One technical point is worth noting: In an ordinal model, the minimum number of necessary threshold constraints also depends on the constraints already in place for the factor loadings (Millsap and Yun-Tein 2004, p. 490). When we released the threshold constraints on the four non-invariant factors but still constrained the factor loadings to be equal across groups, it was not necessary anymore to set two thresholds of *all indicators* to be equal across groups (as suggested by Millsap and Yun-Tein 2004). It was sufficient to set two thresholds of only *one indicator per factor* (instead of two thresholds of all indicators) for the four non-invariant factors to be equal across groups to identify the model (see Appendix).

Table 4: Mean differences of the values for Luxembourg and the Netherlands (compared with Belgium, the reference group)

	Luxembourg		The Netherlands	
	MGCFA Under Assumption of Normality	MGCFA Using Robust WLS	MGCFA Under Assumption of Normality	MGCFA Using Robust WLS
Universalism-benevolence	.00	0.09*	–.12**	–0.18**
Stimulation	0.08*	0.16**	–.02	–0.03
Self-direction	–.01	0.05	0.06	0.05

*$p < .05$; **$p < .01$.

Table 4 displays the means of the values in Luxembourg and the Netherlands. The mean values in Belgium are constrained to zero as Belgium is the reference group (for other identification methods to estimate latent means see Little, Slegers and Card, 2006, and the chapter of Lee, Little and Preacher in this book). There are several significant differences in

the means across countries. In the continuous case (Davidov and Schmidt 2007), people in Luxembourg rate stimulation values as more important (.08) than people in the other countries do and people in the Netherlands rate universalism-benevolence values as less important (-.12) than the others do. The remaining value ratings do not differ across the three countries. Accounting for ordinality leads to similar conclusions in most cases, but mean differences are more pronounced. The higher rating of stimulation in Luxembourg is .16 and the lower rating of universalism-benevolence in the Netherlands is -.18. In addition, the rating of universalism-benevolence is higher in Luxembourg (.09) than in Belgium (.00 as the reference group). However, under the assumption of multivariate normality the latter difference was not significant.

From the analyses that account for ordinality we can conclude that self-direction values are equally important in all three countries. Stimulation values are equally important in Belgium and the Netherlands, but more important in Luxembourg, and the unified value universalism-benevolence is more important in Luxembourg than in either of the other countries. From a substantive point of view, these differences are rather small, reflecting cultural similarities across the Benelux countries. However, we cannot compare means for the other four values because they do not exhibit the necessary level of invariance across countries.

5. Summary and Discussion

In this chapter we tested the comparability of the human values measurement across Belgium, the Netherlands, and Luxembourg using data from the second round (2004-5) of the European Social Survey (ESS). A previous study (Davidov and Schmidt, 2007) established metric invariance across these countries for seven values and scalar invariance for three values, stimulation, self-direction and the unified value universalism-benevolence. That study applied multiple-group confirmatory factor analysis (MGCFA) and assumed a continuous scale with normally distributed responses. However, Lubke and Muthén (2004) have argued that MGCFA is not appropriate for testing for invariance of Likert-type scales. We addressed this criticism by fitting a model for ordinal (ordered-categorical) outcomes to test for invariance.

Previous studies (Welkenhuysen-Gybels and Billiet, 2002; Welkenhuysen-Gybels, 2004; De Beuckelaer and Swinnen's chapter in this volume) have demonstrated, based on simulation studies, that assuming normality and continuity in the case of Likert-type scales generally

does not lead to erroneous conclusions and works well for invariance tests across groups.[7] This chapter assessed whether this conclusion holds with actual survey data on human values as measured in the ESS.

The conclusions with respect to the invariance of the scales were very similar for the model that assumed multivariate normality and the one which accounted for ordinality. The model for ordinal indicators, like the model assuming multivariate normality and continuity, showed that full (or scalar) invariance is present for only three of the seven values – stimulation, self-direction, and the unified value universalism-benevolence. Thus, both models justified comparison across the three countries of the means of these three values. Mean comparisons with the two models yielded fairly similar results. Notably, the mean differences were somewhat more pronounced with the ordinal model and, moreover, only the ordinal model revealed that the mean importance of the unified universalism-benevolence value was significantly higher in Luxembourg. In sum, our results are largely in accordance with simulation studies and justify using MGCFA for Likert scales in the analysis of cross-cultural and longitudinal data. The results also suggest that an ordinal MGCFA for such scales is more powerful to compare latent means.

[7] Though the power of MGCFA is lower than proportional odds modeling (see Welkenhuysen-Gybels and Billiet, 2002, p. 216).

References

Arbuckle, J. L. (2005) *Amos 6.0 User's Guide*. Chicago, IL: SPSS.

Billiet, J. (2003) Cross-cultural Equivalence with structural equation modeling. In: Harkness, J. A., Van de Vijver,F. & Mohler, P. P. (eds.) *Cross-cultural survey methods.* New York: Wiley, 247-264.

Cheung, G. W. & Rensvold, R. B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling* 9: 233-255.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989) Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin* 105: 456-466.

Campbell, D. T. & Fiske, D. W. (1959) Convergent and discriminant validation by the multitrait multimethod matrices. *Psychological Bulletin* 56:81-105.

Chen, F. F. (2007) Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling* 14:464-504.

Cheung, G. W. & Rensvold, R. B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling* 9: 233-55.

Chou, C.-P., Bentler, P. M., & Satorra, A. (1991) Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study. British Journal of Mathematical and Statistical Psychology, 44, 347–357.

Coenders, G. & Saris, W. E. (1995) Categorization and measurement quality. The choice between Pearson and polychoric Correlations. In Saris, W. E. & Münnich, Á. (eds.) *The multitrait-multimethod approach to evaluate measurement instruments.* Budapest: Eötvös University Press, 125-144.

Coenders, G., Satorra, A., & Saris, W. E. (1997) Alternative approaches to structural modeling of ordinal data. A Monte Carlo study. *Structural Equation Modeling* 4:261-282.

Davidov, E. (2008) A cross-country and cross-time comparison of the human values measurements with the second round of the European Social Survey. *Survey Research Methods* 2:33-46.

Davidov, E. & Schmidt, P. (2007) Are values in the Benelux comparable? In Loosveldt, G., Swyngedouw, M. & Cambré, B. (eds.) *Measuring meaningful data in social research.* Leuven, Belgium: Acco, 373-386.

Davidov, E., Schmidt, P., & Schwartz, S. H. (2008) Bringing values back in: Testing the adequacy of the European Social Survey to measure values in 20 Countries. *Public Opinion Quarterly* 72, 420-445.

De Beuckelaer, A. (2005) *Measurement invariance issues in international management research.* Doctoral dissertation, Hasselt University, Diepenbeek, Belgium.

Flora, D. B. & Curran, P. J. (2004) An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* 9:466-491.

Hu, L. & Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* 6, 1-55.

Jöreskog, K. G. (1971) Simultaneous factor analysis in several populations. *Psychometrika* 36:409-426.

Jöreskog, K. G., & Sörbom, D. (1996) *LISREL 8 user's reference guide*. Chicago: Scientific Software International.

Knoppen, D. & Saris, W. (2009) Do we have to combine values in the Schwartz´ Human Values Scale? A comment on the Davidov studies. *Survey Research Methods* 3:91-103.

Little, T. D., Slegers, D. W., & Card, N. A. (2006) A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling*13:59-72.

Lubke, G. H. & Muthén, B. O. (2004) Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling* 11: 514-534.

Marsh, H. W., Hau, K.-T. & Wen, Z. (2004) In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling* 11:320-341.

Millsap, R. E. & J. Yun-Tein (2004) Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research* 39:479-515.

Muthén, B. O. & Asparouhov, T. (2002) *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus.* Mplus Web Notes, No. 4, retrieved from: http://www.statmodel.com/examples/webnote.shtml [23.04.2009]

Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997) Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Technical report, retrieved from: http://www.gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf [18.04.2009]

Muthén, L. K & Muthén, B. O. (1998) *Mplus user's guide*. Los Angeles: Muthén & Muthén.

Muthén, L. K. & Muthén, B. O. (2007) *Mplus user's guid*e. Los Angeles: Muthén & Muthén.

Sagiv, L. & Schwartz, S. H. (1995) Value priorities and readiness for out-group social contact. *Journal of Personality and Social Psychology* 69, 437-448.

Satorra, A. (1992) Asymptotic robust inferences in the analysis of mean and covariance structures. *Sociological methodology 22:249-278.*

Satorra, A. & Bentler, P. M. (1990) Model conditions for asymptotic robustness in the analysis of linear relations. *Computational Statistics and Data Analysis* 10:235–249.

Schwartz, S.H. (1992) Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology* 25: 1-65.

Schwartz, S. H. (1994) Are there universal aspects in the content and structure of values? *Journal of Social Issues* 50:19-45.

Schwartz, S. H. (2005) Basic human values: Their content and structure across countries. In: Tamayo, A. & Porto, J. B. (eds.), Valores e Comportamento nas Organizações [Values and Behavior in Organizations] Petrópolis, Brazil: Vozes, 21-55.

Schwartz, S. H., Melech, G., Lehmann A., Burgess, S., Harris, M., & Owens, V. (2001) Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross Cultural Psychology* 32:519-542.

Sörbom, D. (1974) A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology* 27:229-239.

Sörbom, D. (1978) An alternative to the methodology for analysis of covariance. *Psychometrika* 43:381-396.

Steenkamp, J.-B. E. M. & Baumgartner, H. (1998) Assessing measurement invariance in cross national consumer research. *Journal of Consumer Research* 25:78-90.

Temme, D. (2006) Assessing measurement invariance of ordinal indicators in cross-national research. In Diehl, S. & Terlutter, R. (eds.*) International advertising and communication. Current insights and empirical findings*. Wiesbaden, Germany: GWV, 455- 472.

Vandenberg, R. J. (2002) Towards a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods* 5:139-158.

Vandenberg, R. J. & Lance, C. E. (2000) A review and synthesis of the measurement invariance literature: Suggestions, practices and recommendations for organizational research. *Organizational Research Methods* 3:4-69.

Welkenhuysen-Gybels, J., & Billiet, J. (2002) A comparison of techniques for detecting cross-cultural inequivalence at the item level. *Quality & Quantity* 36:197-218.

Welkenhuysen-Gybels, J. (2004) The performance of some observed and unobserved conditional invariance techniques for the detection of differential item functioning. Quality *& Quantity* 38:681-702.

Appendix

Mplus syntax for the final model with invariance of three values

```
1  title: MGCFA Values ESS Round 2 - Belgium, Luxembourg, Netherlands;
2  data: file is benelux_values.dat;
3  variable: names are ipcrtiv imprich ipeqopt ipshabt impsafe impdiff ipfrule ipudrst
4  ipmodst ipgdtim impfree iphlppl ipsuces ipstrgv ipadvnt ipbhprp iprspot iplylfr
5  impenv imptrad impfun country;
6  categorical are all;
7  grouping is country (1=BE 2=LU 3=NL);
8  missing = all (7-9);
9  Analysis: PARAM = THETA;
10 Model:
11 SD by impfree* ipcrtiv@1;
12 UNBE by ipeqopt* ipudrst@1 impenv iphlppl iplylfr ipadvnt imprich;
13 COTR by ipmodst* imptrad* ipfrule@1 ipbhprp iprspot imprich;
14 SEC by impsafe@1 ipstrgv;
15 POAC by imprich* iprspot* ipshabt@1 ipsuces ipmodst;
16 HE by ipgdtim@1 impfun;
17 ST by impdiff@1 ipadvnt;
18 Model LU:
19 [ipfrule$3* ipfrule$4* ipfrule$5*];
20 [ipbhprp$1* ipbhprp$2* ipbhprp$3* ipbhprp$4* ipbhprp$5*];
21 [imptrad$1* imptrad$2* imptrad$3* imptrad$4* imptrad$5*];
22 [ipmodst$1* ipmodst$2* ipmodst$3* ipmodst$4* ipmodst$5*];
23 [ipstrgv$1* ipstrgv$2* ipstrgv$3* ipstrgv$4* ipstrgv$5*];
24 [impsafe$3* impsafe$4* impsafe$5*];
25 [ipshabt$3* ipshabt$4* ipshabt$5*];
26 [ipsuces$1* ipsuces$2* ipsuces$3* ipsuces$4* ipsuces$5*];
27 [iprspot$1* iprspot$2* iprspot$3* iprspot$4* iprspot$5*];
28 [imprich$1* imprich$2* imprich$3* imprich$4* imprich$5*];
29 [ipgdtim$1* ipgdtim$2* ipgdtim$3* ipgdtim$4* ipgdtim$5*];
30 [impfun$3* impfun$4* impfun$5*];
31 Model NL:
```

```
[ipfrule$3* ipfrule$4* ipfrule$5*];
[ipbhprp$1* ipbhprp$2* ipbhprp$3* ipbhprp$4* ipbhprp$5*];
[imptrad$1* imptrad$2* imptrad$3* imptrad$4* imptrad$5*];
[ipmodst$1* ipmodst$2* ipmodst$3* ipmodst$4* ipmodst$5*];
[ipstrgv$1* ipstrgv$2* ipstrgv$3* ipstrgv$4* ipstrgv$5*];
[impsafe$3* impsafe$4* impsafe$5*];
[ipshabt$3* ipshabt$4* ipshabt$5*];
[ipsuces$1* ipsuces$2* ipsuces$3* ipsuces$4* ipsuces$5*];
[iprspot$1* iprspot$2* iprspot$3* iprspot$4* iprspot$5*];
[imprich$1* imprich$2* imprich$3* imprich$4* imprich$5*];
[ipgdtim$1* ipgdtim$2* ipgdtim$3* ipgdtim$4* ipgdtim$5*];
[impfun$3* impfun$4* impfun$5*];
output: stand; res;
```

Explanation

The syntax in Mplus only needs to specify deviations from the Mplus default. The Mplus default for MGCFAs with ordinal data is:

Parameterization: DELTA

Estimation: WLSMV

Constraints: first factor loading on each “by” statement set to 1, all intercepts set to 0, all factor loadings and thresholds constrained equal across groups

Lines 1-6 Data source, variable names, definition as categorical

Line 7 Multiple groups

Line 8 Indication which categories represent missing values for all variables

Line 9 Theta parameterization

Lines 10-17 Model for all groups, “*” specifies that a parameter should be estimated, “@” fixes a parameter. “SD by impfree* ipcrtiv@1;” overrides the Mplus default to constrain the first factor loading to 1 and takes the second instead.

Lines 18-30 threshold parameters for the four factors that are not invariant across groups are set free in the group “Luxembourg”. Thresholds that are not mentioned explicitly remain constrained across groups automatically (the same applies to the factor loadings),
e.g., “[ipfrule$3* ipfrule$4* ipfrule$5*];” sets the last three thresholds of the item ipfrule free and leaves the first two constrained

Lines 32-43 threshold parameters for the four factors that are not invariant across groups are set free in the group “Netherlands”.

Line 44 Request of standardized estimates and residual statistics in the output.

Paper 2

Two theories on the test bench: Internal and external validity of the theories of Ronald Inglehart and Shalom Schwartz

---

Co-authored with: Wolfgang Jagodzinski, University of Cologne
Peter Schmidt, University of Marburg

Published as: Datler, Georg, Jagodzinski, Wolfgang & Schmidt, Peter (2012) Two theories on the test bench: Internal and external validity of the theories of Ronald Inglehart and Shalom Schwartz. Social Science Research 42, 906-925.

Two theories on the test bench: Internal and external validity of the theories of Ronald Inglehart and Shalom Schwartz

Abstract

In the last decades value research has produced a vast number of theoretical concepts. However, it is unclear how the different value theories relate to each other. This study makes a first step toward a systematic comparison of value theories. It focuses on the individual level of the two approaches that are, at present, probably the most prominent in international research - the theory of basic human values of Shalom Schwartz and the postmodernization theory of Ronald Inglehart. Using data from the World Value Survey and the European Social Survey for West Germany we assess both the internal and the external validity of the two accounts. The results indicate that both value theories have different strengths and weaknesses. Whereas the Inglehart account has lower internal and weaker construct validity, the Schwartz account is somewhat less consistent in its predications. Nevertheless, both value conceptions are able to explain a substantial share of variation in specific attitudes and behavior.

1. Introduction

The empirical research of the last decades has produced an impressive number of different value orientations. Sometimes values are equated with more or less abstract, positively evaluated objects or states: Health, family, work, religion and many other entities are therefore called values. Sometimes values are related to basic human needs, like the needs for security, affiliation, or love. In the classical tradition values are defined as standards such as the values of freedom, equality, justice, or fairness. Apart from these principle disagreements about the concept of values, there are differences with regard to specific values. Two authors may use the same value name but understand and operationalize the underlying value differently or they assign different value names to very similar sets of indicators.

Different value researchers do not completely ignore each other but they quote the studies of others selectively and usually only in those cases where the findings of the other seem to support their own view.[1] Comprehensive studies of the relationships between different value approaches are completely lacking. It is almost certain that problems of discriminant validity would arise if similar values from different theories were included in one and the same study (Jagodzinski 2004). International comparative studies so far do not allow a comprehensive assessment of advanced value theories. It is true that the World Value Survey 2005 also includes ten items of the Portrait Value Questionnaire of Schwartz in addition to the indicators of Inglehart's value dimensions. However, it can already be anticipated that ten items cannot adequately cover the ten broadly defined value orientations of Shalom Schwartz, which is discussed in more detail in sections 3.1 and 3.2. Survey research may be reluctant to include the measurement instruments of different value theories into their questionnaires partly because they do not want to confront the respondents with batteries of similar questions and partly because it would increase the costs of such a survey immensely. Therefore, at the moment it cannot be said whether value research violates Occam's principle and multiplies entities, in this case: values, beyond necessity. It is very likely that it does but no one can presently prove this.

In order to overcome the present situation, this paper attempts to systematically compare two very prominent value theories, the theory of basic human values of Shalom Schwartz and the postmodernization theory of Ronald Inglehart (e.g. 1977). Both authors present two-level theories, which distinguish between macro-level cultural values and individual-level value orientations. It is true, the focus of Inglehart's (e.g. 1977) research has recently shifted to such an extent to the macro-level that the micro-level component of his theory can be overlooked. As the postmaterialism theory is only rudimentarily integrated into the new, more encompassing approach, one may gain the impression that we actually deal with two theories, a micro-level theory of postmaterialism and a macro-level theory of self-expression values. This is not the view of Ron Inglehart, however. Even his publications on macro-level cultural change persistently emphasize that cultural change is the result of micro-level value change (see, e.g., Inglehart, 1997; Norris and Inglehart, 2004; Inglehart and Welzel, 2005). The

[1] Hofstede (2001), in his presentation of individualism/collectivism refers to Triandis as well as to Inglehart's post-materialism. Similarly, Inglehart and Welzel (2005) identify autonomy as the common theme of individualism (Triandis) and self-direction (Schwartz) and self-expression values, but they do not analyze these relationships in detail.

analysis of Inglehart and Baker (2000) further shows that cultural values and individual-level value orientations are operationalized with the same set of indicators. Due to space limitations, we have to confine ourselves exclusively to the key concepts of the individual-level value orientations in both approaches, which for the sake of brevity will be simply called values.

A comparison of two value theories should, first of all, investigate the internal validity of the measurement. Recent methodological studies on the measurement instruments of the ESS give important insights into this field, particularly also into problems of measurement equivalence, but they investigate only rudimentarily the predictive power of the underlying concepts. This is largely consistent with the strategy of Schwartz and his colleagues who mainly concentrated on the internal structure and validity of the values and only sparsely examined the relations between values and external variables. As long as this part of the theory remains less developed, however, the theory is of limited interest for the nonexperimental social sciences, which have always seen the main attraction of value theories in their promise to explain a broad range of beliefs, attitudes, and behaviors by a limited number of values. A comparison of value theories can, therefore, not be based on internal validity alone (Jagodzinski and Manabe, 2009; Opp and Wippler, 1990). Relationships with external variables, which either predict values or are predicted by values, are at least as important.

As both theories relate values to a set of common external variables, the strength and signs of these relationships will be the second criterion, which for the sake of brevity is called external validity of the theory. A theory is externally valid if all relationships have the theoretically predicted signs and the explained variance in all dependent variables is high.[2] Though the predictive power of the values is in the focus of interest, the paper will also investigate the effect of selected exogenous variables on values.

Besides internal and external validity, the parsimony is used as a third standard of comparison. If two value theories have more or less the same explanatory power, the one with fewer values is more parsimonious and, therefore, superior to the other. So we have three criteria which we apply step by step to the two value theories. Before we do this, we very

[2] Clearly, the external validity is estimated under the assumption that the model is correctly specified. As long as there is no empirical evidence, however, that low external validity is a result of spurious non-correlations, they indicate problems of the examined theory.

briefly discuss communalities and differences in the theories of Ronald Inglehart (1977) and Shalom Schwartz (1992, 1994). The internal and external validity are examined in Section 3. As the study has to rely on two separate surveys, we use the European Social Survey (ESS) 2004 for measuring the values of Schwartz and the World Value Survey (WVS) 2005 for measuring the values of Inglehart. Needless to say, the external validity can only be assessed with regard to those external variables which are at least similarly measured in both surveys. Results are summarized and discussed in the last Section

2. The Two Value Theories – Similarities and Differences

Space limitations do not allow a comprehensive discussion of the two theories. The values of both theories will be very briefly described and compared in Subsection 2.1. The basic features of the measurement models are examined next (Subsection 2.2). The last Subsection discusses the relationship between values and a subset of external variables, which are similarly measured in ESS 2004 and WVS 2005. These relationships are summarized in a set of hypotheses (Subsection 2.3).

2.1 The Value Concepts

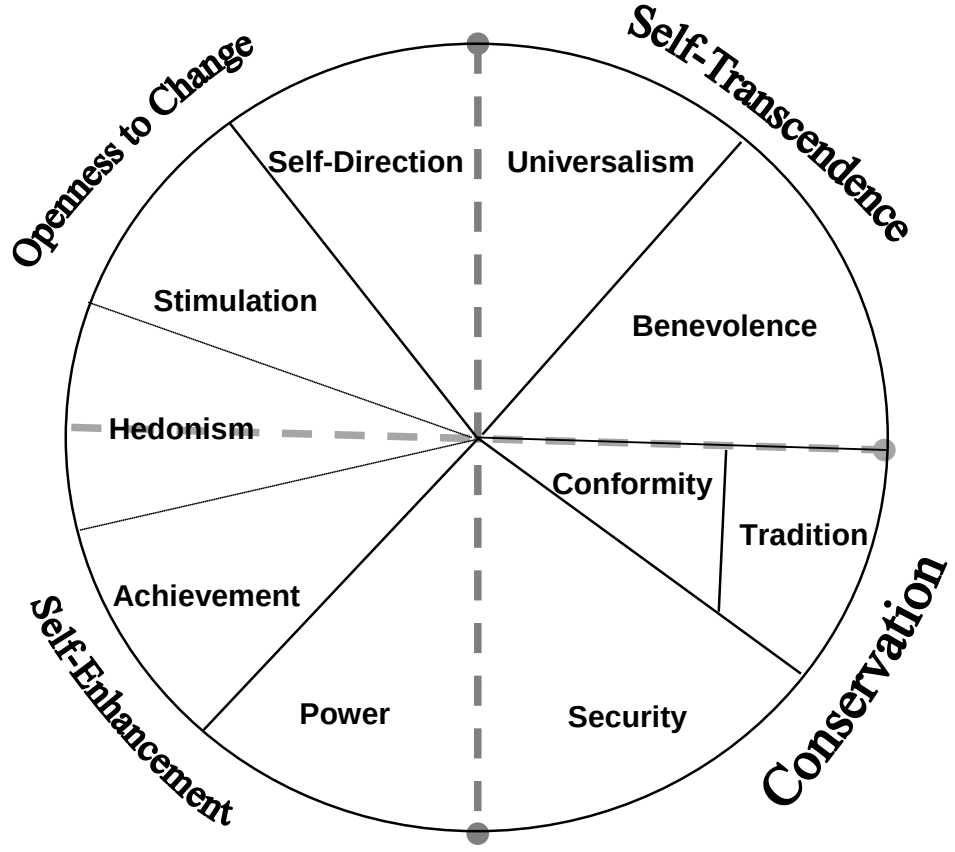
a) Schwartz. The value theory of Schwartz proposes **ten** basic values that are intended to include all the main values recognized across cultures in the world (for a new extension to more than ten values, see Schwartz and Vecchione, 2011): power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity, and security. These values and the motivations behind them are presented in Table 1. They are derived from three universal requirements of human beings: needs as biological organisms, coordinated social interaction, and survival and welfare needs of groups (Schwartz, 1992, 1994, 2007). Furthermore, Figure 1 displays the relations between the values. Two important features may be observed. First, some values oppose each other whereas others are congruent with each other. Pursuing tradition and conformity may be congruent. However, pursuing tradition values may oppose following self-direction values. Second, the values are grouped behind two higher order dimensions. The first higher order dimension contrasts self-transcendence with self-enhancement values. The second higher order dimension opposes openness to change with conservation values.

Table 1: Schwartz' ten basic human value types and the motivations behind them

Value type	Motivational emphasis
Power	Social status and prestige, control or dominance over people and resources
Achievement	Personal success through demonstrating competence according to social standards
Hedonism	Pleasure and sensuous gratification for oneself
Stimulation	Excitement, novelty and challenge in life
Self-direction	Independent thought and action - choosing, creating and exploring
Universalism	Understanding, appreciation, tolerance and protection for the welfare of all people and for nature
Benevolence	Preservation and enhancement of the welfare of people with whom one has frequent personal contact
Tradition	Respect, commitment and acceptance of the customs and ideas that traditional culture or religion provide
Conformity	Restraint of actions, inclinations and impulses likely to upset or harm others and violate social expectations or norms
Security	Safety, harmony and stability of society, or relationships, and of self

Source: Sagiv and Schwartz (1995: 438)

Figure 1: The Relations between the Ten Values



b) Inglehart. Inglehart (1977) relates values to human needs. The first version of his approach, the so-called postmaterialism theory, reduced Maslow's hierarchy to two basic needs, which are called materialistic and postmaterialistic. They were conceptualized as poles of a unidimensional value continuum. Materialism becomes manifest in a preference for material and physical security, postmaterialism in the emphasis on higher needs like freedom, participation, self-expression, or beauty. The second version defines this dimension more broadly as survival/self-expression dimension and includes interpersonal trust, happiness, and a liberal sexual morality as further indicators. Moreover, a second dimension is added which contrasts so-called traditional and secular-rational societies. Both dimensions are imbedded in a theory of value change, which describes modernization as a two-phase process. While traditional are replaced by secular-rational values in the process of industrialization, the self-expression values become predominant during the transition from the industrial to the postindustrial or advanced societies.

Discussions on the dimensionality of materialism and postmaterialism, particularly Flanagan's (1987) distinction between an economic and an authoritarian-liberal dimension, may have had an impact on Inglehart's revisions (Inglehart, 1997; Inglehart and Baker, 2000), but altogether the new values rather look more like inductive generalizations from empirical findings rather than theoretically derived concepts.[3]

c) Similarities. Schwartz specifies ten values, Inglehart four or two – depending on whether the poles or the dimensions are counted as values: Nevertheless, there are similarities between the concepts. *Security* (lower right segment of Figure 1) corresponds to survival needs, and stimulation and self-direction in the upper segment correspond to self-expression quite well. The match between a traditional orientation (Inglehart) and tradition (Schwartz) is obvious. Achievement in the lower left segment could be interpreted as a secular-rational orientation because it was an essential outcome of the modernization process. The two value dimensions of Inglehart can be plausibly located in the value space of Schwartz though the orthogonality may not be preserved.

[3] Inglehart and Welzel (2005) have further elaborated the theoretical model of value change, but left the measurement model unchanged.

2.2 The Measurement Models

a) Inglehart. The first version of Inglehart's theory only included two alternative measurement models for the Materialism/Postmaterialism (MPM)-Dimension. The larger model consists of twelve items[4], the smaller one, the so-called 4-item MPM-index, is derived from naming the two top priorities among four political goals. As only two questions – each with four priorities – are required, the 4-item MPM-index is extremely parsimonious and presumably the most widely used value measurement in international surveys. The discussion of the index would fill a whole methodological textbook of insightful criticism and intriguing counterarguments. Even the most ardent critics cannot deny, however, that the index fits Kluckhohn's (1951) famous definition of a value remarkably well: It measures conceptions of the desirable [features of a society][5], and it is shown that these conceptions determine attitudes and behavior in various domains.

The 4-item MPM-index indeed includes a relatively small number of political preferences. The objection that it remains a measure of political preferences cannot be fully denied. The measurement instrument of the revised theory covers a much broader range of topics but it also has become very heterogeneous, including feelings, emotions, and reported behavior. The MPM-index along with questions on happiness, homosexuality, interpersonal trust, and on political petitions are used for measuring survival/self-expression values. Religiosity, condemnation of abortion, deference to authority, obedience, and identification with the nation characterize traditionalists and distinguish them from secular-rational persons. No doubts that there are other and probably better measures of the two value dimensions, but Inglehart being interested in long-term value change, has decided for the ten indicators in Table 2 (Inglehart and Baker, 2000) because they are available in all European Values and World Values Surveys since 1981. As a consequence, the secular-rational pole of the value continuum is not positively defined by indicators of a rational orientation but only negatively determined as the absence of traditional orientations.

---

[4] De Graaf *et al*. (1989) report reasonably high factor loadings of the 12 items, after correction for ipsative measurement has been performed.

[5] Respondents have to choose whether higher priority should be given to freedom and participation or to the maintenance of order and economic stability

Table 2: The Measurement of the Inglehart Values in the WVS 2005

Traditional vs. Secular-Rational Values		
1	Importance of God	How important is God in your life? Please use this scale to indicate. 10 means "very important" and 1 means "not at all important. (***10-point scale***)
2	Autonomy Index[a]	Here is a list of [10] qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Please choose up to five! (*Points in brackets are added or subtracted if the respective item is mentioned)* …, Independence (+1), …, Determination, perseverance(+1), …, Religious faith (-1), Obedience (-1) (***5-point scale***)
3	Abortion[a]	Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card: (1=never justifiable, …, 10=always justifiable) ….., Abortion … (***10-point scale***)
4	National Pride[a]	How proud are you to be [**French**]? (1=very proud, .., , 5=not at all proud). (***5-point scale***)
5	Respect for Authority	I'm going to read out a list of various changes in our way of life that might take place in the near future. Please tell me for each one, if it were to happen, whether you think it would be a good thing, a bad thing, or don't you mind? (1= Good, 2=Don't mind, 3=Bad): …, Respect for Authority, … (***3-point scale***)
Survival vs. Self-expression Values		
1	Materialism-Postmaterialism Index[b]	People sometimes talk about what the aims of this country should be for the next ten years. On this card are listed some of the goals which different people would give top priority. Would you please say which one of these you, yourself, consider the most important? 1) Maintaining order in the nation 2) Giving people more say in important government decisions 3) Fighting rising prices 4) Protecting freedom of speech And which would be the next most important? (Scale derived from the two top priorities: 1= Materialists: aims 1) and 3) mentioned; 3= Postmaterialists: aims 2) and 4) mentioned; 2= other combinations mentioned. (***3-point scale)***
2	Happiness[b]	Taking all things together, would you say you are: 1=Very happy, 2=Rather happy, 3=Not very happy, 4=Not at all happy. (***4-point scale***)
3	Homosexuality[b]	Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card. (1=never justifiable, …, 10=always justifiable): ….., Homosexuality (***10-point scale***)
4	Sign a petition	Now I'd like you to look at this card. I'm going to read out some forms of political action that people can take, and I'd like you to tell me, for each one, whether you have done any of these things, whether you might do it or would never under any circumstances do it: ….; Signing a petition. (3-point scale: 1=have done; 2=would do; 3=would never do. (***3-point scale***)
5	Interpersonal Trust	Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? (Code one answer): 1=Most people can be trusted, 2=Need to be very careful. (***2-point scale***)

[a] Low scores indicate traditional values.
[b] Low scores indicate survival values.

In contrast to Rokeach (1973) or Schwartz (1992), Inglehart avoids the use of generic terms in the questions on values for the reason that specific items are better understood by the respondents and may be less affected by framing effects. On the other hand, items about political preferences are theoretically only indirectly related to values. This is even truer for the other value indicators such as questions on self-reported behavior, interpersonal trust, authoritarianism, or national identity which often are used as indicators for other theoretical constructs like interpersonal trust, authoritarianism and national identity. The factor analytic models which Inglehart and others (cf. Inglehart and Baker, 2000; Inglehart and Welzel,

2005) apply presuppose a direct effect of the latent values on these indicators which clearly is inconsistent with the results of social psychology (see Eagly and Chaiken, 1993; Fishbein and Aizen, 2010). The effect of a value on behavior, for instance, is mediated among others by the evaluation of objects and behavioral intentions, etc. (Fishbein and Aizen, 2010). Furthermore, the reasoned action approach (Fishbein and Aizen, 2010) assumes causal relations among variables that Inglehart treats as indicators of the value variables, that is, the axiom of local independence is violated.[6]

All in all, the reader gains the impression that the indicators of postmodern values have been chosen in such a way as to maximize the relationships with external variables. Inglehart was much less concerned about a coherent expansion of the original individual-level value model as the new indicator *Happiness* shows. We know, from previous studies, that happiness does not correlate with postmaterialism: "Subjective well-being is a condition, not a value, and not correlated with Postmaterialism at the individual level" (Inglehart, 1997, p. 87). We infer that the MPM-index and the *Happiness* item should not be combined as multiple indicators in a measurement model for individual-level postmodern values because sufficiently large indicator correlations are a minimum requirement of high reliability. From the sentence that immediately follows: "But high levels of subjective well-being are a key element in the cultural syndrome called Postmodernism." (Inglehart, 1997, p. 87), we might further conclude that Inglehart wants to use different indicators for the macro-level cultural syndromes and the individual level values – a decision which as such could not be criticized. Inglehart and Baker (2000), however, take exactly the same indicators for measuring macro- and micro-level values. Two consequences can already be anticipated before any empirical analysis is carried out: The use of both indicators will deteriorate the reliability and internal validity of the measurement model and, at the same time, increase the correlation of the survival/self-expression index with well-being. The latter result has a tautological flavor because it is near-at-hand that an index including happiness as a component will predict well-being quite well. We will return to that problem in a later Section.

[6] The MPM Index as a general political attitude has an indirect causal (mediated via intention) effect on "signing a petition" (behavior).

b) Schwartz. To avoid indirect indicators, Schwartz strictly distinguishes between the measurement of values and the measurement of attitudes (Schwartz, 2007). He tries to tackle the confounding of values and attitudes by proposing an instrument of broad and basic motivations relevant to various attitudes and behaviors in different domains in life (Schwartz, 2007). This strategy, however, also has its price. Expressions like: "everybody in the world be treated equally", for instance, can be interpreted in the sense of equal starting conditions or in a strict egalitarian sense which would include equal pay for all human beings, etc. People may agree with the first idea but not with the second. Depending on how the question is framed we have to expect different answers.[7] These framing effects may not only produce random measurement error but also result in halo effects and other kinds of systematic distortions. Unfortunately, there has been so far no systematic comparison of Ingelhart's and Schwartz' measurement instruments using cognitive interviews (Willis, 2005), which would give us better insight into the measurement quality of these items. The indicators of Schwartz are listed in Table 3.

[7] It is likely that a respondent will be more positive toward this question if previous questions referred to discrimination of minorities and more negative if previous questions referred to equal pay for industrious and idle workers.

Table 3: The Human Values Scale in the ESS 2004

Value	Item # (according to its order in the ESS questionnaire) and Wording (Male Version)
Self-Direction (SD)	1. Thinking up new ideas and being creative is important to him. He likes to do things in his own original way (ipcrtiv). 11. It is important to him to make his own decisions about what he does. He likes to be free to plan and not depend on others (impfree).
Universalism (UN)	3. He thinks it is important that every person in the world be treated equally. He believes everyone should have equal opportunities in life (ipeqopt). 8. It is important to him to listen to people who are different from him. Even when he disagrees with them, he still wants to understand them (ipudrst). 19. He strongly believes that people should care for nature. Looking after the environment is important to him (impenv).
Benevolence (BE)	12. It is very important to him to help the people around him. He wants to care for their well-being (iphlppl). 18. It is important to him to be loyal to his friends. He wants to devote himself to people close to him (iplylfr).
Tradition (TR)	9. It is important to him to be humble and modest. He tries not to draw attention to himself (ipmodst). 20. Tradition is important to him. He tries to follow the customs handed down by his religion or his family (imptrad).
Conformity (CO)	7. He believes that people should do what they're told. He thinks people should follow rules at all times, even when no one is watching (ipfrule). 16. It is important to him always to behave properly. He wants to avoid doing anything people would say is wrong (ipbhprp).
Security (SEC)	5. It is important to him to live in secure surroundings. He avoids anything that might endanger his safety (impsafe). 14. It is important to him that the government insures his safety against all threats. He wants the state to be strong so it can defend its citizens (ipstrgv).
Power (PO)	2. It is important to him to be rich. He wants to have a lot of money and expensive things (imprich). 17. It is important to him to get respect from others. He wants people to do what he says (iprspot).
Achievement (AC)	4. It is important to him to show his abilities. He wants people to admire what he does (ipshabt). 13. Being very successful is important to him. He hopes people will recognize his achievements (ipsuces).
Hedonism (HE)	10. Having a good time is important to him. He likes to “spoil” himself (ipgdtim). 21. He seeks every chance he can to have fun. It is important to him to do things that give him pleasure (impfun).
Stimulation (ST)	6. He likes surprises and is always looking for new things to do. He thinks it is important to do lots of different things in life (impdiff). 15. He looks for adventures and likes to take risks. He wants to have an exciting life (ipadvnt).

2.3 The Hypotheses

While Schwartz at least implicitly emphasizes value stability, Inglehart presents a dynamic theory of value change, though he shares the assumption that individual-level values are by and large stable in adulthood. Inglehart was, from the early seventies on, intrigued by the observation that the values of younger generations markedly differed from the values of the parent generations in the West. It was not a small shift from giving a higher priority to a given value to giving a higher priority instead to a neighboring value in the Schwartz circle, say

from conformity to security, but rather a shift from one side to the opposite side in Figure 1: While the parents still strive more for security and material goods compared to younger generations, the younger generations emphasize self-actualization and the abolishment of old hierarchical structures more strongly than their parents. The seminal book *'The Silent Revolution'* from 1977 tries to explain the value differences between post-war older generations and younger generations as a change from materialistic to postmaterialistic values. Inspired by Maslow's need hierarchy and the principle of relative potency, Inglehart systematically elaborated the idea that economic, social, and political security has a tremendous impact on human life. People who grew up and lived in a secure environment differ in political attitudes and behavior, gender and family orientations, work preferences, religious orientations, environmental concern, interpersonal trust, and many other issues from those who have to fight for their subsistence and are threatened by turmoil, wars, and disorder.

Along these lines he has developed a rich set of highly differentiated hypotheses – most of which appear already in the first version of his theory. It has already been shown by example of *Happiness*, however, that the relationship between postmaterialist and postmodern values is not always clear. It is sometimes doubtful whether and how the hypotheses of the former theory can be translated into hypotheses for the new values. As long as there is no convincing argument to the contrary, however, we infer, from the fact that the MPM-index is seen as a key indicator of the new survival/self-expression dimension, that postmodern individualists feel, think, and behave like postmaterialists. The relationships between the traditional/secular-rational value dimension and external variables are even less developed. When not otherwise stated, we tentatively infer the respective hypotheses from the basic ideas of the theory.

As far as the hypotheses of Schwartz are concerned, we strictly will confine ourselves to those hypotheses which have been explicitly stated in the literature. Sometimes opposite effects are stated for opposite values in Figure 1 - but can we generally assume that opposite values display opposite relationships with external variables? Is a positive effect of age on conservation, for example, necessarily paralleled by a negative effect of age on openness to change, and a positive influence of benevolence on immigration by a negative effect of self-enhancement on the same attitude? If the value space were truly two-dimensional, this would be the case. Two (not perfectly correlated) values would be sufficient, however, for extracting the two orthogonal dimensions. All other values could be calculated as a linear combination of these two dimensions or of the two values. Accordingly, "opposite values" would be linearly dependent, and their effects could not be estimated. Actually the values of Schwartz

are located in a space of higher order. The projection into the two-dimensional space represents the relationships among the values approximately but is far from perfect. As a consequence, we cannot always predict from the relationship of a first value with an external variable the sign of the effect of the opposite value (see Costner and Leik, 1964). This introduces additional complexities into theory construction and testing as will be seen later on.

Below we present a brief summary of those hypotheses that can be tested with the ESS 2004 and the WVS 2005 data. This is a fairly severe restriction because both surveys do not include many comparable external variables. Most of the items refer to the socio-political domain which may slightly bias the results in favor of the postmodernization theory which often is called a political theory. As long as Schwartz does not systematically specify the influence of values on attitude and behavior, however, the size of the bias remains unknown.

There is a further qualification: Frequently, the surveys include a general variable such as the left-right self-placement scale, but an appropriate test of a theory would require a finer distinction, for instance, a differentiation between the economic and the libertarian/authoritarian meaning of left and right. If this is not possible, we can only estimate the presumably weaker effect on the general left-right scale and in this regard probably underestimate the predictive power.

Bearing these qualifications in mind, we first discuss the relationships of values with the exogenous variables age, gender, and education, and then move to factors, which are dependent on values.

2.3.1 The Impact of Social-Structural Variables on Values

Age and Generation. To Inglehart, generational differences are a function of economic and technological change. Younger generations differ from older insofar as they grow up in a different environment. Inglehart has never completely excluded life-cycle effects (see, e.g., Inglehart, 1981, 1990, 1997; Inglehart and Welzel, 2005), but he interprets age differences in the first place as a sign of generational change. Large economic changes, as they occurred in the West after World War II, also imply the already mentioned large differences between the older materialistic and the younger postmaterialistic generations. The trend could also reverse if the advanced society would experience a long-lasting period of economic decline.

In the later theory of postmodernization, value change occurs on two dimensions: Those who grow up in very poor, traditional societies internalize traditional and survival values; those

who grow up in affluent, postindustrial societies, place an emphasis on rational-secular and self-actualization values in adulthood. In the transition from a traditional to an industrial society, generational differences along the traditional/secular-rational axis should be most pronounced while in the transition from the latter to the postindustrial society, differences on the survival/self-expression dimension should be greater.

It is not necessary to elaborate entirely the model of generational change because we investigate values in a single society at a single point in time. It is sufficient to state the implication of the model for the effect of age on the values: As Germany has experienced a fast economic development after World War II, younger people should be more rational-secular than older people, and they should also place more emphasis on self-actualization values. Age effects cannot be separated from generation effects in a cross-sectional analysis. If the latter should be superimposed by life-cycle effects (Jagodzinski, 1983) the relationship would only be strengthened. Accordingly, we expect a negative effect of age on both value dimensions.

Schwartz (2006, 2007) also refers to both cohort effects due to differential socialization contexts and individual life-cycle effects. In the derivation of concrete hypotheses for age, Schwartz follows the logic of life-cycle effects. Older people become more committed to their habits and develop stronger ties to their social networks, resulting in fewer changes and challenges, less openness for such changes, and stronger reliance on conservative values (Glen, 1974). Therefore, older people are expected to score higher on conservation values (tradition, conformity, security) and lower on openness to change values (self-direction, stimulation) (Schwartz, 2006, 2007; Meuleman et al., in press). Furthermore, as individuals become with age less preoccupied with themselves and more with others, self-transcendence is expected to increase and self-enhancement to decrease (Schwartz, 2006; Veroff et al., 1984).

Education. If education enhances exposure to new experiences, different people, and alternative ways of reasoning, more educated individuals should score higher on self-enhancement, openness to change, and self-transcendence values, and lower on conservation values (Meuleman et al., in press; Schwartz, 2006). According to Inglehart, respondent's education to some extent reflects the indoctrination of values in school, but to a larger extent the economic and social conditions during the formative years: Children of highly educated

families usually grow up in a secure environment[8]. Thus, the higher the education, the more likely secular-rational and self-expression values have been internalized (Inglehart, 1990): Education, in other words has a negative effect on traditional and survival values.

Gender. Inglehart explicitly states no value differences between men and women. Schwartz, by contrast, follows theories of gender differences in assuming that men attribute higher importance to instrumental values such as power or achievement, as well as stimulation, hedonism, or self-direction values, and lower importance to benevolence and universalism values (Schwartz, 2006, 2007; Schwartz and Rubel, 2005; Schwartz and Rubel-Lifschitz, 2009; Meuleman et al., in press).

2.3.2 The Impact of Values

Left-Right Self-Placement. Postmaterialists as early as in the seventies placed themselves on the left of the left-right continuum, because they aimed at political change, opposed the old, middle to right-wing political elites, and held leftist political and social attitudes. Inglehart emphasized very early that the meaning of left and right undergoes a change in advanced societies, from the economic left-right distinction to a new conflict, which centers around self-expression values, and becomes manifest in issues of environmental protection, gender roles, or political participation. The meaning of the left-right dimension is also dependent on the locations of parties in a given country. Thus, if the political parties hold positions along the economic left-right dimension, this understanding of left and right will also dominate in the public. For the West German public the notions *left* and *right* have a mixture of meanings. They are still associated with the economic cleavage where a leftist position is egalitarian. Emphasis on gender equality, environmental protection, and liberal moral attitudes are also seen as left positions. Right by contrast, is not only associated with conservatism, law and order, and hierarchy but also with church and religiosity. Accordingly, traditional as well as survival values should be positively correlated with a right position on the scale.

In the framework of Shalom Schwartz, individuals scoring high on universalism and benevolence values (self-transcendent dimension); who are open to change, or who focus on the welfare of the others are expected to be more strongly related to the political left, whereas individuals scoring high on conservation and self-enhancement values are expected to be

[8] Education of the parents would better reflect the economic conditions in childhood and youth. The explanatory power is slightly underestimated if respondent's education is used as a predictor.

more strongly related to the political right. These expectations correspond to the meaning of left and right in contemporary liberal democracies (Piurko et al., forthcoming; see also Schwartz, 2006; Caprara et al., 2006).

Political Interest. With regard to political interest and political activism the situation is slightly more complicated. In many Western countries, party identification and voter turnout rates have declined during the last decades so that it seems obvious to expect a decline of political interest, too. Inglehart, however, has argued that postmaterialists are more interested in politics (Inglehart, 1997, p. 308). Self-actualization indeed requires political engagement in democratic societies so that political interest can be expected to increase with self-expression values. Whether the same holds for secular-rational values is doubtful. During the postwar era, churches in Western Europe supported social and political integration. A good Christian had to vote in elections and be interested in politics as well. Accordingly, the relationship between secular-rational values and political interest may be much weaker or even be reversed.

Because of the similarity between postmaterialism and universalism and maybe benevolence values as well, one could also infer that individuals who place high importance on self-transcendence values may be more interested and involved in politics. Indeed, politics often involves issues that are of key concern for universalistic ideologies such as immigration policies, social welfare, or environment. Security, conformity, and self-direction values have also turned out to be of central importance in the formation of political values of voters in various contexts (see, e.g., Barnea, 2003; Schwartz, 2006; Caprara et al., 2006).

Political Activism. More than thirty years ago Inglehart predicted a decline of elite-directed and an increase of elite-challenging political behavior (Inglehart, 1977, pp. 317-321). The former comprises voting and traditional forms of party support, in particular party identification; the latter the participation in boycotts and protest demonstrations or so-called unconventional protest behavior. The prediction was based on two central variables, cognitive mobilization and postmaterialism. The former would increase as a consequence of rising formal education and the improvement of political skills, the latter as a consequence of rising economic prosperity. Elite-challenging political behavior should still be positively affected by education[9] and self-expression values or, vice versa, negatively affected by survival values.

[9] An index of cognitive mobilization would certainly be a better predictor than education.

As elite-challenging political activism remains the least wide-spread among traditional people, traditional values should also have a negative effect.

One may expect universalism and benevolence values, which promote social justice and environmental care--important topics in politics--- to predict stronger political activism (Schwartz, 2006). Since political activism may be risky and involve excitement and formation of independent and maybe deviating opinions, security and conformity are expected to display the most negative association with political activism (Schwartz, 2006, 2007).

Attitudes toward Immigrants. Inglehart (1977, p. 320) discusses early on the link between a cosmopolitan orientation, cognitive mobilization, and postmaterialism, whose basic mechanism is elaborated more clearly in later publications. Feelings of insecurity are the main source of parochialism and xenophobia (Inglehart, 1997). Existential threats foster ingroup/outgroup thinking. People seek safety in closely knit networks. Accordingly, traditional as well as survival values are negatively related to openness toward immigrants.

As far as the Schwartz values are concerned several studies (Schwartz, 2006, 2007; Davidov et al., 2008b; Davidov and Meuleman, 2012) link universalism and benevolence values to a positive, and conservation and conformity to a negative attitude toward immigrants. The latter, negative relation can be expected because conservatives tend to perceive immigrants as a threat to the preservation of norms, customs, and established religions, or, in short, the stability of a society. By contrast, the motivational goals of self-transcendence values (especially universalism), such as understanding, appreciation, tolerance, and protection for the welfare of people and for nature, are promoted by the arrival of immigrants (e.g., Sagiv and Schwartz, 1995)

Life Satisfaction. We have already mentioned the fact that life satisfaction may correlate with the self-actualization index for the simple reason that *Happiness* is a component of the index. However, there is also a theoretical reason for expecting a positive relation between self-expression values and happiness or life satisfaction. Individualists have more influence on the definition of their political, social, and private goals and are less directed by others. They, therefore, also have better chances of realizing their goals, can attribute the success to themselves, and be more proud of their actions. As a consequence, they reach a higher level of satisfaction (see Jagodzinski, 2010a for a more comprehensive analysis). By way of contrast, survival values should negatively affect life satisfaction.

In the value system of Schwartz one may consider different views on how values may affect life satisfaction. One view suggests that successfully realizing any of one's values may

increase one's well-being. According to this view, when people believe that some values or goals they see as important are blocked, their life satisfaction suffers (Schwartz and Melech, 2000). Other authors suggest that only certain values may have an impact on well-being. Jensen and Bergin (1988) and Strupp (1980) identified 'healthy' values (self-direction, benevolence, universalism, but also achievement and stimulation) and 'unhealthy' ones (conformity, tradition, security, power) (see Sagiv and Schwartz, 2000). They argue that *realizing* 'healthy' values should enhance happiness. Yet the questionnaire does not measure the realization of values but only their importance and one cannot infer from the importance of values that they are realized. Bilsky and Schwartz (1994) hypothesized that values representing growth (self-direction, universalism, benevolence, achievement, and stimulation) are emphasized by individuals who successfully realize these values in their lives and display higher levels of well-being. In contrast, when these values cannot be realized, individuals tend to suffer from lower levels of well-being and increasingly emphasize conformity, security and power values (but also tradition) compensating for deprivation (Bilsky and Schwartz, 1994; Sagiv and Schwartz, 2000; see also Schwartz et al., 2000). Although hedonism points toward personal pleasure as a goal, Sagiv and Schwatz (2000, p. 182) maintain that one cannot derive a theoretical relation of the importance of hedonistic values and life satisfaction. Empirically, they do not find a significant correlation between hedonistic value orientation and subjective well-being.

Church Attendance. Traditional people do not only hold strong beliefs, but they also participate regularly in religious services. Therefore, a positive correlation between religiosity and traditional values can be expected. The same holds for survival values because religion also gives security and orientation in an insecure environment. Applying the same logic to the values of Schwartz, one can expect a positive impact of traditional individuals on the frequency of religious participation. Hedonists, by contrast, should attend religious services less frequently (Schwartz and Huismans, 1995) because, after all, religion signifies preferring transcendental concerns over materialistic ones. Benevolence and conformity also include some degree of selflessness and are expected to be positively related to religiosity. Since religion is also related to preserving the status quo and increasing certainty in life, one may expect religiosity to be positively linked with tradition, conformity, and security values, and negatively with stimulation and self-direction values, that reflect openness to change (Schwartz and Huismans, 1995).

Attitudes toward Gender Equality. One of the core elements of self-expression values is the emancipation from authority. Human beings are free and equal. Self-expression values are,

therefore, positively and survival values negatively correlated with gender equality. Traditional values should have a similar negative effect because the differential treatment of men and women is firmly anchored in the ideology and religion of traditional societies. Following a similar logic, a preference for conservation values in the Schwartz theory is expected to predict a rejection of gender equality. A universalistic orientation toward humankind and a benevolent orientation to the concrete others (theoretically, men and women) promote a positive stance on gender equality. As self-enhancement and stimulation emphasize individual freedom and openness to new experiences, they should lead to a preference for gender equality as well. Table 4 summarizes the hypotheses.

Table 4: The Hypotheses

		AGE	EDUCATION	GENDER : Female		LEFT-RIGHT SCALE	POLITICAL INTEREST	POLITICAL ACTIVISM	OPENNESS TO IMMIGRANTS	LIFE SATISFACTION	CHURCH ATTENDANCE	GENDER EQUALITY
Traditional (vs. secular-rational) Values		+	-			+		-	-		+	-
Survival (vs. self-expression)Values		+	-			+	-	-	-	-		-
SELF-TRANSCENDENCE	Benevolence	+	+	+		-	+	+	+	+	+	+
	Universalism	+	+	+		-	+	+	+	+		+
OPENNESS TO CHANGE	Self-Direction	-	+			-				+	-	+
	Stimulation	-	+			-				+	-	+
Hedonism											-	+
SELF-ENHANCEMENT	Achievement	-	+	-		+				+		-
	Power	-	+	-		+				-		-
CONSERVATION	Security	+	-			+	+	-	-	-	+	-
	Conformity	+	-			+	+	-	-	-	+	-
	Tradition	+	-			+	+	-	-	-	+	-

Notes: "+" positive relation expected; "- "negative relation expected; empty cells: no relation expected/ theoretical relation unspecified.

Next we turn to the empirical part. In this section we will describe the datasets used and the measurement characteristics of both value theories (internal validity) and present the relations of values in both theories with external variables (external validity and parsimony).

3. Empirical Analyses

3.1 General Problems and Limitations

Before the analysis can be carried out, two important decisions have to be made. The first is favorable to Inglehart. Jagodzinski has carried out the same factor analysis as Inglehart and Baker (2000) with the data of the WVS 2005 and has shown that only the West German factor pattern was in line with the former results (see below, section 3.2 a). Therefore, we confine our analysis to West Germany (N = 1,851 in the ESS, N = 988 in the WVS).[10]

The second decision is in favor of Schwartz. Faced with the question of whether to carry out the analysis exclusively with the WVS or to test the theory of Shalom Schwartz with the ESS, we have opted for the latter alternative for several reasons. Firstly, 21 indicators (ESS) are usually better than ten (WVS), because there are at least two indicators available to measure each of the ten values (see Bollen, 1989). Had we presented the results for the ten-indicator models of the WVS, it is highly likely that objections would have been raised that these few indicators do not cover the values of Schwartz adequately. Secondly, 21 indicators measure the underlying values more reliably than ten if the indicator correlations are sufficiently high. We will show later on that this rule even applies if five generalized values instead of the ten basic Schwartz values are specified. Thirdly, we will almost certainly estimate larger effects and higher predictive power of the Schwartz values by using the ESS because some of the dependent orientations are estimated more reliably.[11]It finally has to be mentioned that Inglehart's value dimensions have to be operationalized as simple additive indices. This method does not eliminate the random measurement error completely and tends to underestimate the strength of the relationships among latent variables (Bollen, 1989). The measurement of the Schwartz values is also not optimal but the number of items in the ESS is at least sufficiently large to measure each of the five generalized Schwartz values by at least

[10] The response rate of the German ESS 2004 was 52.6% (ESS2 – 2004, Data Documentation Report, Edition 3.3). The response rate of the German WVS 2005 was 46.3% (own calculation, based on "Technical Information" available on the WVS website, formula RR1, AAPOR guidelines, see http://www.aapor.org). For detailed information on the survey programs and data access, see www.worldvaluessurvey.org and www.europeansocialsurvey.org.

[11] It is one of the strategic goals of the ESS to develop reliable and valid measurement instruments for a set of relatively homogenous European countries. To the extent that this program is successful, the variables in the ESS will have lower random measurement error than variables in a worldwide survey and, as a consequence, yield higher standardized effects and higher R-squares (see Bollen, 1989 for the general argument). Accordingly, lower measurement error in the dependent variables may positively affect the predictive power estimates for the Schwartz values.

two indicators. As a consequence, a correction of measurement error can be performed with SEM (structural equation model) programs like AMOS which usually results in stronger relationships among the latent variables compared with an analysis which only makes use of summative indices.[12] It has to be added that Inglehart as well as Schwartz and many of their colleagues are working on the improvements of the measurement models for the values. [13] Both theories are still under construction (see below). The future measurement instruments will certainly overcome some limitations of the existing approaches.

3.2 Internal Validity

(a) Inglehart

The use of factor scores or summative indices is quite widespread in comparative research. In the former case, factor loadings serve as typical criteria of reliability, in the latter case, Cronbach's alpha. Inglehart and Baker (2000) report factor loadings and calculate factor scores. A reliable measurement instrument of stable values should also produce a more or less stable factor pattern over time. Jagodzinski (2010b) has, therefore, tried to replicate the factor pattern of the micro-level analysis of Inglehart and Baker (2000) with the WVS 2005 data. In his reanalysis, he used a very soft criterion of similarity or reproducibility: The factor pattern is successfully replicated if (1) the indicators display the same pattern of main loadings as in Inglehart and Baker (2000, p. 24), (2) the main loadings are above .4, and (3) always higher in magnitude than the cross (side) loadings.[14] For that purpose, principal component analyses (PCA) were carried out in all countries and in the pooled data set. Missing data were deleted pairwise, and the number of factors was fixed at two. The solution was subjected to a varimax rotation. Contrary to the expectations, the factor pattern of Inglehart and Baker did not emerge

---

[12] Suppose that the indicators x and y of two latent variables display a correlation of .4 and that 50 percent of the variance in x and y is random measurement error. In this case we will estimate a correlation of .8 among the latent variables.

[13] Welzel (2010), for instance, tries to solve some of the measurement problems by using formative indicators and modified values. A discussion of these models is beyond the scope of the present article but see Diamantopoulos et al. (2008) for a critical discussion of the potentials and the limitations of formative measurements. In particular, correction for measurement error is not as straightforward in formative measurement models as in reflective ones.

Current methodological studies on the Schwartz values propose a refined measurement with more indicators and/or more latent variables (Knoppen and Saris, 2009b; Beierlein et al., 2012).

[14] More rigorous criteria, which are becoming state of the art, were formulated by Meredith (1993).

in the pooled dataset or in any country (see Jagodzinski, 2010b for some of the results). In a separate analysis of the German data the resulting model for West Germany met the three criteria (see the left part of Table 5) while the models for East Germany (middle part of Table 5) and all-Germany (right part of Table 5) did not. The factors for West Germany are estimated in such a way that high positive scores indicate a preference for tradition and for survival. For indicating the direction of the value dimension we simplify the notation: instead of *Traditional/Secular-Rational values* we speak of *Traditional Values*, and instead of *Survival/ Self-Expression values* of *Survival Values.*

Table 5: Replication of Inglehart and Baker (2000) with the German Data of WVS 2005, Principal Component Analysis, Standardized Factor Loadings

	West		East		Total	
Secular-Rational Values Indicators	**TRAD**	**SURV**	**?**	**?**	**?**	**?**
V192 God is very important	**.660**	-.191	**.494**	-.452	.507	**-.518**
Y003 Obedience over Independence	**.488**	.273	**.494**	.009	**.536**	.029
V204 Abortion is never justifiable	**.595**	.347	**.726**	.002	**.720**	.018
V209 R is very proud of nationality	**.476**	.060	**.380**	-.301	**.418**	-.109
V78 Respect for authority is good	**.585**	.066	**.434**	-.012	**.464**	.079
Self-Expression Values Indicators						
Y002 R is materialist	.273	**.497**	**.469**	.333	**.431**	.399
V10 R is unhappy	-.299	**.600**	-.019	**.619**	-.065	**.600**
V96 R would never sign a petition	.236	**.479**	**.392**	.377	.371	**.406**
V202 Homosexuality is never justifiable	.475	**.574**	**.718**	.160	**.685**	.267
V23 Need to be careful with people	-.016	**.576**	.105	**.642**	.103	**.634**
Explained Variance (%)	**37.78%**		**36.18%**		**37.06%**	

Notes: Forced 2-factor solution, Varimax rotation; pairwise deletion of missing data, N=953.
TRAD stands for the traditional/secular-rational and SURV for the survival/self-expression dimension. Traditionalists have a high positive score on TRAD and security-oriented people a high positive score on SURV. A question mark on the top of the last columns of the table indicates that the labels TRAD and SURV are inappropriate; R = respondent; high factor loadings are indicated in bold.

Although the principal component analysis for West Germany distinguished the two theoretical dimensions, this result departs from a simple structure. In particular, the cross-loadings of the items on abortion and on homosexuality are substantial.[15] The former was conceptualized by Inglehart and Baker (2000) as a measure of the traditional/secular-rational value orientation, the latter as a measure of the survival/self-expression value orientation. However, both indicators are highly correlated among each other, and that is one of the

[15] From a methodological point of view, principal axis factoring (PAF) is more suitable to discover latent dimensions that are measured by observed indicators because PAF, in contrast to PCA, includes the assumption of measurement error. We employ PCA to replicate the original study of Inglehart and Baker (2000).

obstacles to the specification of a parsimonious confirmatory factor analysis (CFA) model. We tried out a larger number of specifications but never arrived at a satisfactory result.[16]

Inglehart investigated measurement problems of materialism and postmaterialism in the 1970s and 1980s (see De Graaf et al., 1989; Inglehart, 1985). It is understandable, though regrettable, that the development of postmodernization theory was not paralleled by a similar elaboration of suitable measurement instruments for his new value concepts. We follow Inglehart in the subsequent analysis and work with similar additive indices and factor scores.

(b) Schwartz

Schwartz and several other researchers have invested a considerable amount of work in the elaboration of the underlying measurement model. It is now widely agreed, however, that even the 21 Portrait Values Questionnaire (PVQ) items in the ESS are not sufficient to measure all ten values adequately. Whereas studies using the full 40-item version of the PVQ (Steinmetz et al., 2009; Schmidt et al., 2007; Beckers et al., 2012) could identify all ten theoretical values in the empirical analysis, studies with the shortened ESS version of the PVQ report high correlations between items intended to measure adjacent but different values. Knoppen and Saris (2009a) have questioned the face validity of the indicators of universalism, tradition, and power and have proposed as a possible strategy to work for the time being with those seven values which are reliably measured. In addition, Schwartz et al. (in press) have enlarged the model to 19 values in a new cross-national study. Other authors have collapsed adjacent values (Bezonsky et al., 2011, Davidov et al., 2008 b, Duriez et al., 2005, Verkusalo et al., 2009). Davidov et al. (2008a) and Davidov (2008, 2010), for instance, identify maximally seven distinct value constructs across most of the ESS countries by using all ESS items in a multiple group confirmatory factor analysis (MGCFA). They model the indicators of three pairs of adjacent values as measuring three new latent variables: the indicators of benevolence (BE) and universalism (UN) are now specified as indicators of a new latent variable UNBE, those of achievement (AC) and power (PO) as indicators of POAC, and those of tradition (TR) and conformity (CO) as indicators of TRCO. With some modifications (see Davidov et al., 2008a, pp. 434f.), this model achieves satisfactory fit and can be accepted in terms of internal validity. From a broader perspective, these different approaches focus on the same values because they use the original indicators of Schwartz and

[16] Results of the CFA models for the Inglehart items can be obtained from the authors upon request.

specify the measurement models in such a way that the resulting latent variables remain in close vicinity semantically to the corresponding values of Schwartz. The SEM methodology is much stricter, however. It is not sufficient to use the same indicators and value names in different studies. Rather, metric equivalence requires that also the unstandardized factor loadings do not differ (Meredith, 1993); for the minimum condition of partial metric invariance at least two unstandardized factor loadings per construct have to be equal (Byrne and Van de Vijver, 2010). Different studies typically impose different constraints on the measurement model, estimate different unstandardized factor loadings, and therefore also generate different values. This is also true of our study. It was not our intention to create a new set of values but we could not reproduce the values of former studies.

It is a property of the circumplex structure that indicators of adjacent values sometimes correlate as high with each other as with the other indicator(s) of the same value, and the correlation between adjacent values, therefore, is also very high. In the model with seven values, self-direction and stimulation as well security and conformity/tradition correlate above .80. Using all seven value constructs in a multiple regression results in extraordinary large standard errors which indicate problems of multicollinearity.[17]

To ensure that the value constructs can be used in regression models and to preserve the desirable proportions of CFA models with correction for measurement error instead of using additive indices, we have to collapse further indicators to therefore create new and broader concepts. This strategy is not solely data-driven but distinguishes those four values which Schwartz conceptualizes on a higher level: ”Openness to change” (OPEN) – measured in our analysis by indicators of self-direction and stimulation, “self-transcendence” (SELF-TRANS) - measured by indicators of universalism and benevolence , “conservation” (CONSERV) – measured by the indicators of conformity, tradition, and security, and finally “self-enhancement” (SELF-ENH) – measured by the indicators of power and achievement. While Schwartz introduces these values as second-order concepts, we specify them as new first-order latent variables. From the original values only hedonism (HE) remains and is situated between openness to change and self-enhancement.

[17] This is one of the reasons why structural equation modeling (SEM) with the 7 values as predictors of external variables more often than not produces non-admissible solutions (generally for this problem of multicollinearity in SEM see Marsh et al., 2004).

Table 6: Confirmatory Factor Analysis, Standardized Factor Loadings

	SELF-TRANS		CONSERV		SELF-ENH		HE		OPEN	
	WVS	ESS	WVS	ESS	WVS	ESS	WVS	ESS	WVS	ESS
ipeqopt		.459								
Ipudrst		.615								
Impenv	.724	.534								
Iphlppl	.674	.648								
Iplylfr		.552								
ipmodst				.444		-.333				
Imptrad	.379		.278	.513						
Ipfrule				.542						
Ipbhprp			.516	.696	.316					
Impsafe	-.120		.653	.632						
Ipstrgv				.512						
Imprich	-.486	-.271			.734	.566				
Iprspot				.218		.434				
Ipshabt						.659				
Ipsuces					.625	.722				
Ipgdtim							1.00*	.759		
Impfun								.726		
Impdiff										.680
Ipadvnt		-.416							.653	.796
Ipcrtiv	.364								.417	.527
Impfree										.512

* fixed to 1 for identification

Notes: ESS 2004, West German sample, pairwise deletion of missing data, N = 1,832; WVS 2005, West German sample, pairwise deletion of missing data, N=953; gray fields: ex post modifications (cross-loadings); Bold items are included in ESS and WVS, (ipbhprp, ipgdtim, imprich are identical in the two surveys, the other items differ in wording), for item wording in the ESS see Table 3; SELF-TRANS = Self-Transcendence, CONSERV = Conservation, SELF-ENH = Self-Enhancement, HE = Hedonism, OPEN = Openness to Change;
Model fit ESS: Chi2 = 961.4 with df = 175, RMSEA = 0.050 (p close = .591), CFI = .912
Model fit WVS: Chi2 = 160.6, df = 22, CFI = .918, RMSEA = .081, p-close = 0.000

As the number of values has been reduced in any case, the question may arise of whether the ten indicators of the WVS are not sufficient for a suitable measurement model. This had the obvious advantage that one and the same data set could be used and independent and dependent variables in our tests would be exactly the same. The two measurement models presented in Table 6 allow us to answer the question. The standardized loadings of the ten WVS indicators are always displayed in the first column below the respective values, the loadings of the 21 ESS indicators in the second. Cross loadings which are not consistent with the theoretical expectations are shaded gray. A brief comparison tells us that the ESS model fits the theoretical expectations much better than the WVS model. While the ten WVS indicators display five theoretically unexpected cross loadings, the 21 ESS indicators display only four. The WVS indicator “Important to be rich” displays a very high negative loading on

SELF-TRANS though one can hardly say that there is a direct impact of self-transition on the desire to be rich. Finally, tradition (Imptrad), which should be a key indicator of CONSERV in the WVS, actually has a higher loading on SELF-TRANS (.379 vs. .278).[18]

These results argue strongly in favor of the ESS solution. The correlations among the latent variables for this model are displayed in Table 7. To achieve an acceptable fit ($Chi^2 = 961.4$ with df = 175, RMSEA = 0.050, CFI = .912), four cross-loadings have to be allowed. All cross-loadings are substantially smaller than the respective target loadings. Three of them are negative in sign. In these instances, the indicator is in a sense more distant from the side factor than the other indicators of the main factor. The only positive cross-loading that had to be allowed relates an indicator designed to measure self-enhancement (in fact, the subdimension of power) to the theoretically adjacent dimension of conservation. Apart from these deviations, however, the five latent variables are measured fairly well.

Table 7: Correlations of the Latent Factors in the Schwartz Model

	SELF-TRANS	CONSERV	SELF-ENH	HE	OPEN
SELF-TRANS	*.221*				
CONSERV	.382	*.531*			
SELF-ENH	.061	.080	*.498*		
HE	.171	-.156	.531	*.727*	
OPEN	.426	-.223	.620	.778	*.410*

Notes: *Variances in the diagonal*; ESS 2004, West German sample.

Notice that the average size of the correlations between the five values is much lower than those found between the seven value constructs identified by Davidov et al (2008a). Therefore, this conceptualization is more suitable for multiple regression models in which different values are simultaneously used as predictors. So we arrive at a viable, but certainly suboptimal, solution.

[18] Cronbach's α supports the choice of ESS, too. If the standardized indicators with loadings >.4 in magnitude are summed up to an index of the respective values, the ESS value index always surpasses the WVS value index in terms of reliability. The alphas are .636 (WVS) vs. .670 (ESS) for SELF-TRANS, .418 (WVS) vs. .596 (ESS) for OPENESS, .533 (WVS) vs. .674 (ESS) for SELF-ENH, and .485 (WVS) vs. .694 (ESS) for SELF-TRANS. While α=.710 for hedonism in ESS, no alpha can be calculated in the single WVS indicator. It is true, the reported figures are in general not very impressive. Yet while the estimates for the ESS scores in most of the cases come close to .7, which is often seen as the lower bound for a suitable scale, the α reliabilities of the WVS indices always remain below that margin. We do not want to discuss the pros and cons of Cronbach's α here because we will not use sum scores for measuring the values of Schwartz. We take the results as a further indication, however, that also the latent variables in SEM models would suffer from reliability problems if the ten items of the WVS were used for measuring even only five latent values.

3.3 External Validity

The external validity of both value theories is assessed with regard to their relation to a selection of external variables. The selection was guided by two criteria: (1) Theoretical relevance for both value theories (see section 2.3). (2) (Almost) equal or similar measurements in the WVS and ESS (for item wordings and response scales see Appendix, Table A1, and Table A2).

(a) Inglehart

Table 8: Relationship of Survival and Traditional Values with External Variables, Standardized Regression Coefficients

Values as Dependent Variables								
						SOCIODEMO		
	R^2					Age	Fem	Educ
Sum Scores								
Traditional values (TRAD)	.156					.348	n.s.	-.112
Survival values (SURV)	.126					.225	n.s.	-.220
Factor Scores								
Traditional values (TRAD)	.120					.319	n.s.	-.071
Survival values (SURV)	.092					.177	n.s.	-.201
Values as Independent Variables								
		R^2		VALUES		SOCIODEMO		
	TOT	Δ VAL	DEMO	TRAD	SURV	Age	Fem	Educ
Sum Scores								
Left-right scale	.104	.059	.045	.213	.101	n.s	-.147	-.096
Political interest	.156	.047	.109	.154	-.228	.169	-.158	.236
Political activism	.272	.167	.105	-.104	-.395	n.s.	n.s.	.144
Openness to immigration	.154	.081	.073	n.s.	-.290	n.s.	.116	.119
Life satisfaction	.178	.157	.021	.178	-.443	n.s.	n.s.	n.s.
Church attendance	.328	.210	.118	.523	-.133	.180	.118	.079
Gender equality	.145	.041	.104	-.136	-.134	-.101	.128	.157
Factor Scores								
Left-right scale	.105	.060	.045	.249	.103	n.s.	-.149	-.104
Political interest	.140	.031	.109	.078	-.163	.175	-.155	.242
Political activism	.244	.139	.105	-.266	-.318	n.s.	n.s.	.160
Openness to immigration	.150	.077	.073	-.093	-.286	n.s.	.116	.124
Life satisfaction	.249	.228	.021	.178	-.451	n.s.	n.s.	n.s.
Church attendance	.334	.216	.118	.490	$-.041^{n.s.}$	.183	.120	.077
Gender equality	.147	.043	.104	-.183	-.142	-.095	.128	.160

Notes: WVS 2005, West German sample, pairwise deletion of missing values; n.s. = non significant ($p > .05$); method of factor score estimation: regression; SOCIODEMO = sociodemographic variables; Fem = female; Educ = education; TOT = total explained variance; Δ VAL= additional explained variance when values are added as independent variables into the model; DEMO = the explained variance when only sociodemographic variables are included as independent variables in the model; TRAD= traditional values; SURV = survival values.

Table 8 reports the effects of sociodemographic variables on survival and traditional values in the upper part and the effect of these values on external variables when controlling for demographic variables in the lower part. As hypothesized, traditional values and survival values are related positively to age and negatively to education. In general, using values as

predictors of attitudes and behavior confirms the theoretically expected relationships (see section 2.3). The only exception is the prediction of openness to immigration. A negative effect of both traditional and survival values seemed plausible, but only the effect of survival values is significant.

Although the effects of the Inglehart values are substantial on all external variables, there are noticeable differences in the predictive power of the different external variables. Survival values and traditional values can explain more than 20% of the variance in church attendance and life satisfaction and around 15% of the variance of political activism over and beyond sociodemographic variables. At the same time, the effect size is much lower on political interest, self placement on the left-right scale, and attitudes toward gender equality with around 5% of additional explained variance for each. For openness to immigration, the effect size is in the medium range with an 8% change in R^2. Using factor scores (compared to simple sum scores) does not consistently result in higher explained variance. The R^2 for life satisfaction differs most. The model with unweighted sum scores explains only 18% of the variance and the model with factor scores, in contrast, explains more than 25%.

(b) Schwartz

Taking advantage of the measurement model presented above, all analyses for the Schwartz values were conducted in a multivariate model in an SEM framework with latent variables.[19]

Table 9 reports the effects of sociodemographic variables on Schwartz values in the upper part and the effect of these values on external variables when controlling for sociodemographic variables in the lower part. The effects of sociodemographic variables on self-transcendence and openness to change are in line with the theoretical expectations (see section 2.3): Age has a positive effect on self-transcendence and a negative effect on openness to change, education leads to the enhancement of both, and women, on average, score higher on self-transcendence than men. For hedonism, no effects of sociodemographic background variables were expected, but the analysis revealed a negative effect of age and gender on hedonism (with men scoring higher on hedonism). Self-enhancement, in line with the theoretical expectation, is higher for men, increasing with education and declining with age.

[19] Computations were carried using maximum likelihood estimation in AMOS 18. Detailed information on the model specification can be obtained from the authors upon request.

The hypotheses for conservation are confirmed as well: Age has a positive effect, education a negative one. In addition, women tend to be slightly more conservative than men.

Table 9: Relationship of the Schwartz Values with External Values, Standardized Regression Coefficients

Values as Dependent Variables			
		SOCIODEMO	
	R^2	Age	Fem	Educ
Self-Transcendence (SELF-TRANS)	.085	.070	.224	.203
Conservation (CONSERV)	.152	.358	.087	-.152
Self-Enhancement (SELF-ENH)	.171	-.361	-.131	.165
Hedonism (HE)	.123	-.310	-.151	n.s.
Openness to Change (OPEN)	.149	-.360	n.s.	.176

Values as Independent Variables											
	R^2			VALUES					SOCIODEMO		
	TOT	ΔVAL	DEMO	SELF-TRANS	CON-SERV	SELF-ENH	HE	OPEN	Age	Fem	Educ
Left-right scale	.159	.108	.051	-.468	.460	-.177	n.s.	.356	.143	-.037	.069
Political interest	.262	.078	.184	.307	-.212	n.s.	n.s.	n.s.	.211	-.211	.227
Political activism	.200	.118	.082	.366	-.328	n.s.	-.086	n.s.	n.s.	n.s.	.161
Openness to immigration	.222	.130	.092	.494	-.274	n.s.	n.s.	-.180	-.257	n.s.	.093
Life satisfaction	.079	.068	.011	.244	n.s.	n.s.	.301	-.210	n.s.	n.s.	.082
Church attendance	.139	.085	.054	-.198	.399	.178	.269	.390	.063	.069	n.s.
Gender equality	.273	.096	.177	.348	-.442	.172	n.s.	-.206	-.211	.163	.149

Notes: ESS 2004 (Edition 3.2, released 02-02-2011), West German sample; pairwise deletion of missing values; n.s. = non significant ($p > .05$); SOCIODEMO = sociodemographic variables, Fem = female; Educ = education; TOT = total explained variance; ΔVAL= additional explained variance when values are added as independent variables into the model; DEMO = the explained variance when only sociodemographic variables are included as independent variables in the model.

Evaluating the estimated relations of Schwartz values with external variables in light of the theoretical hypotheses we developed gives a mixed picture. As expected, self-transcendence values lead to a more leftist political orientation, and conservation values lead to a more rightist view. The hypothesis that self-enhancement values are a precursor of rightist orientation is clearly rejected for the West German sample, in fact the data point to the opposite. In addition, an unexpected positive effect of openness to change values on rightist views was found. The results for the prediction of "political interest" only partially match the theoretical expectations: Whereas the positive effect of self-transcendence is confirmed, the hypothesized positive effect of self-enhancement is not found and an unexpected negative effect of conservation values on political interest is revealed in the multivariate analyses. With respect to political activism, the expected positive effect of self-transcendence values and the expected negative effect of conservation values are supported by the data. For openness toward immigration, the observed negative effect of conservation values and the positive effect of self-transcendence values correspond to the theoretical expectations, whereas the analyses revealed an additional negative effect of openness to change values. In the prediction

of life satisfaction, only the positive effect of self-transcendence values is confirmed by our analysis. The expected negative effect of conservation values is no longer significant in the multivariate model, and openness to change values show a negative effect where we expected a positive effect. Contrary to the reasoning and the results of previous studies (Bilsky and Schwartz, 1994), we do find a relatively strong positive effect of hedonistic value orientation on life satisfaction. The hypotheses on church attendance only find partial support as well: Respondents high on conservation values go to church more often. But contrary to the theoretical expectation, the effect of openness to change values on church attendance is positive and the effect of self-transcendence is negative. In addition, we observe negative effects of self-enhancement and hedonism values. With respect to attitudes toward gender equality, the expected negative effect of conservation values and the positive effect of self-transcendence values are supported by the data. However, contrary to the hypotheses, the effect of openness to change values is negative in a multivariate model.

Some of these findings seem to unsettle widely-held premises That openness to change is associated with a right position, for instance, contradicts a persistent finding in Western democracies. It is also a near-at-hand conclusion from the value circle because openness to change is located opposite of conservation. Our results call all these considerations and findings into question because openness to change has a relatively strong positive impact (=.356) on the left-right scale. One might try to theoretically explain this result by differentiating between politics and other domains. Openness to change in the private world, it might be argued, must not to be mixed up with a preference for political change. Politically conservative people can be self-directed and open to change in the private world. To them a stable political system is a precondition for being open. The explanation, however, ultimately undermines the idea that values are overarching and not domain-specific orientations. Before we follow this line of reasoning, however, we better look for alternative, in particular also statistical reasons which may have produced the result. We have already pointed out in section 2.3 that opposite values need not have opposite effects if the value space actually has more than two dimensions. This, however, is the case as the correlations in Table 7 show. Nowhere do we detect correlations of -1 for opposite values. The values are also not evenly distributed in the two-dimensional space. Three of them – self-enhancement, hedonism, and openness – cluster closely together with inter-correlations above .5 while conservation is

located farther away from this cluster.[20] As a consequence, the effects of all five values on the dependent variables are identified. Some of the correlations in Table 7 are fairly high but not high enough to cause multicollinearity problems.

We are confronted with another type of problem which is known from the discussion of suppressor effects: The sign of the effect of X on Y may change when a third variable Z is additionally included in the analysis. Openness to change displays the theoretically expected negative correlation with the left-right scale – not very strong but significant (r = -.10). In the multivariate analysis this negative correlation is decomposed into a direct effect and a larger number of indirect relationships. Let us focus for the moment on only one additional variable, namely, self-transcendence, and let us further assume that the impact of -.468 on the left-right scale in Table 7 were a correlation. The correlation between the two values is .442[21] and the indirect relationship between openness to change and the left-right scale via self-transcendence were -.468x.442 ≈-0.20. The effect of openness to change on the left-right scale were about +0.13 (≈ (-.10+.20)/(1-.442^2)).[22] The indirect effect via self-transcendence would have transformed the negative correlation into a positive effect. Actually, the correlation between self-transcendence and left-right-self-placement is lower (= -.130) but in principle the same logic applies: Not just one but several indirect relations of openness to change with the left-right self-placement contribute to a negative relationship – for example, the paths via self-enhancement and conservation, leaving the additional paths via the demographic variables aside.[23] As the sum of all indirect relationships is much below -.10, they have to be counterbalanced, so to speak, by a fairly strong positive effect of openness to change on the dependent variable in order to reproduce a weak negative correlation. The inclusion of other values destroys a theoretically plausible relationship in that it turns a theoretically expected

---

[20] Furthermore, self-enhancement and openness, for example, should correlate more strongly with their immediate neighbor, hedonism, than with each other. Actually, self-enhancement correlates higher with openness (r = .620) than with hedonism (r = .531).

[21] It is in the model with all exogenous variables marginally higher than in Table 7 (0.426). Self-enhancement also displays slightly different correlations with self-transcendence (.083 instead of .061) and with conservation (.069 instead of .080). All other correlations differ only at the third digit. OPEN with SELF-TRANS: .426 / .442

[22] In the three-variable case, the standardized regression coefficient is: $ß_{LO} = (r_{OL} - r_{LT} * r_{OT})/(1 - r_{OT}^2)$ where r_{OL} and r_{TL} are the correlations of the left-right self-placement scale (L) with openness to change (O) and self-transcendence (T), respectively, and r_{OT} is the correlation between the latter two values.

[23] Openness to change displays a positive correlation with the first two values and a negative with the third (see Table 7). Self-transcendence and self-enhancement have a negative effect on the left-right scale and conservation has a positive effect (see Table 9).

negative correlation into a theoretically unexpected positive effect. The precise conditions can be easily stated for models with only two independent variables, while they are somewhat more complicated for larger models. Relatively small changes in the correlation coefficients, in any case, can result in fairly large changes of the regression effects. And these small changes already can occur if indicators are confounded. Two of the indicators of self-transcendence, for instance, are relatively close to leftist positions in Germany: Equal opportunities (ipeqopt) is a major concern of the traditional left and looking after the environment (impenv) a major concern of the new left. The measurement may have a slight political bias and inflate the correlation between self-transcendence and the left-right scale. As a further consequence it can also distort the effect of openness to change.

There are other theoretically unexpected effects which are less counterintuitive. The negative effect of age on hedonism, for example, is consistent with the results of studies on generational change which have shown that younger German generations have become more materialistic and hedonistic than the preceding so-called postmaterialistic generations. We hesitate to interpret these theoretically more plausible findings as conclusive evidence as long as the doubts about our structural model persist. We have to wait and see whether our findings can be replicated and confirmed by other studies.

3.4 Summary: Inglehart vs. Schwartz

On balance there are three dependent variables for which all expected theoretical relations are confirmed (political interest, political activism, attitudes toward immigrants), but for the other five dependent variables, the equations contain not only unexpected effects but empirical findings that clearly contradict the theoretical expectations. A count of confirmed vs. rejected hypotheses yields 100% (6 of 6) correct predictions for the relation of Inglehart values and sociodemographic variables vs. 80% for the Schwartz values (12 of 15). The theoretically expected relationships with independent variables were confirmed for 71% (10 of 14) in the Inglehart case and for 57% (20 of 35) in the Schwartz case. Thus, in terms of correct predications, postmodernization theory performs better than the theory of basic human values.

The effects of Schwartz values are substantial for all external variables. The predictive power is lowest for life satisfaction ($\Delta R^2 = 7\%$) and highest for openness to immigration ($\Delta R^2 = 13\%$). For all other external variables chosen, Schwartz values can explain around 10% of variance above and beyond the sociodemographic variables.

The theoretically derived hypotheses for Inglehart values were supported by the data very well. In the case of Schwartz values, the results were mixed. Some hypotheses could be

confirmed, others were rejected, and still other relationships that were not anticipated became apparent in the data.

The predictive power of the Schwartz values can partly be attributed to theoretically unexpected effects and may have been increased by the correction of attenuation which we could perform in the covariance structure model. In order to retain an acceptable fit of this model we had to create a new set of values – values which are similar to those of other studies, because they are based on the same set of indicators, but which are equivalent neither with the ten nor with the seven nor with other sets of values which have been identified in other studies so far. We also had to accept theoretically unexpected cross-loadings. Furthermore, several of our predictor variables had theoretically unexpected effects, this is particularly true for openness to change. Had we based our predictions in Table 9 only on those values which actually have the theoretically expected effect denoted in Table 4, the R-squares in Table 9 would almost certainly have decreased considerably. Though it may be misleading to qualify R^2 under these conditions as a measure of predictive power, our exploratory strategy has at least the advantage that we can also detect those effects of values which theoretically were not expected.

The partial success of postmodernization theory in explaining "church attendance", "political activism", and "life satisfaction" is related to a different problem. Technically speaking, Inglehart includes indicators in the measurement instruments of value orientations which are strongly related to or confounded with external variables. If, for example, the item *Importance of God* is used as an indicator of traditional values, it is not surprising at all that the value index correlates highly, say, with church attendance. This is sometimes called a tautological explanation. We do not use this expression, first because it is not tautological in a strict sense, second because the appeal to a tautology in empirical research is often conceived as an insult, and third because the problem can be easily explained without bitter remarks: We would strongly overestimate the influence of the value orientation if we do not specify a direct or indirect relationship[24] between the confounded indicator and the respective external variable. Typical symptoms of the problem are that a value indicator correlates higher with external variables than with other value indicators or that the best value indicator does not consistently

[24] In the example of church attendance we may either say that importance of God has a direct impact on church attendance because strong believers also participate more often (direct effect), or we may specify church attendance and importance of God as indicators of a further latent variable, i.e., religiosity.

display the highest correlation with external variables. The latter problem occurs in the Schwartz model, too, but in Inglehart's measurement model these distortions are much stronger. "Happiness" and "signing a petition" are indicators of survival vs. self-expression values which in turn should explain life satisfaction and political activism. Happiness and life satisfaction, however, can be seen as indicators of well-being and "signing a petition" an indicator of political activism. Even though we did not include "signing a petition" into the index calculation of political activism, the problem remains that the value indicator is directly linked to political activism and not only indirectly via the value. Therefore, some of the effects are probably overestimated. Nevertheless, the encompassing theoretical framework and the large set of empirically confirmed hypotheses is the strength of Inglehart's values theory. It is the result of a long and fruitful research that has always tried to understand and to explain value change as well as cultural change. Schwartz, on the other hand, has invested time and energy in the elaboration of a coherent value system and its operationalizations. He conceptualizes values as trans-situational goals or general guidelines that impact on more specific attitudes. That they can explain attitudes toward various objects has been shown in our analysis.

4. Summary and Conclusion

Theories of values and value change help us to understand differences between individuals and cultures. Yet social scientists have identified so many values during the last decades that the question arises of whether we really need them all. From this perspective, the two, presently, most prominent micro-level value theories have been investigated - the value theories proposed by Shalom Schwartz and Ronald Inglehart.

The World Value Survey 2005 includes all items for measuring the values put forth by Inglehart but only ten indicators for the values put forth by Schwartz. We have shown that the latter set of items is much too small to measure ten or five broadly defined values reliably. The European Social Survey 2004 includes 21 items of Schwartz' Portrait Value Questionnaire and is better suited for that purpose but it contains none of Inglehart's value indicators. Therefore, we had to use two different surveys for our comparisons, the World Value Survey 2005 and the European Social Survey ESS 2004. Comparisons of the internal and external validity of both theories were only possible across surveys. The analyses of Inglehart's values were performed with the WVS data, the analyses of the Schwartz' values with the ESS data.

Space constraints as well as the lack of comparable indicators in both datasets did not allow the analysis of more than seven external variables which are predicted by both theories. Therefore, near at hand is the objection that we did not choose the correct ones thereby leading us to not correctly assess the strengths of the theories. It could also be argued that we misunderstood the theory and derived false hypotheses. However, and to the best of our knowledge, we have focused on external relationships in the present analyses that are either well established or at least under discussion.

Though the corresponding external variables always have a similar meaning in both surveys, most of them are operationalized in different ways. The better measurement quality of the ESS items as well the application of different statistical models may have may have shifted the odds slightly in favor of the theory of Shalom Schwartz. Moreover, our analysis was confined to West Germany – the only region in which we could approximately replicate Inglehart and Baker's (2000) factor pattern using the data from the WVS 2005. A final objection could be the fact that our models were much too simple as far as the relationships with external variables are concerned, in particular, the relationships between a value orientation and reported behavior. In the present analysis we followed the common practice in which direct effects of values are specified on attitude and behavior.

There are no doubts that the present study has room for improvement. Nevertheless, our analysis provides, by and large, an adequate description of the present state of the two value theories. Owing to Ronald Inglehart we have an important theory of value change which helps us to understand the ongoing changes in attitudes and behavior in the west and east. His theory makes relatively clear predictions about what will happen in countries like China or India if the economic growth should continue. During the last decades he has elaborated the relationship between values and external variables and, thereby, met the expectations of sociologists and political scientists, who are at least as much interested in their predictive power as in the values themselves. This also became apparent in the empirical analysis, where we could predict the effects in most of the cases correctly. The micro-level effects are significant but markedly smaller than the corresponding macro-level effects, and they are partially inflated by confounded indicators – indicators which are directly related to external variables.

Inglehart's theory is persuasive because it is parsimonious and informative; Schwartz' theory is persuasive because its measurement instruments are theoretically derived. The idea of Schwartz that there are more than four values conceptualized as the poles of two orthogonal

dimensions is probably shared by many other scientists. It is also plausible that some values are more closely related to each other than others. It is a challenging task, however, to reconcile such a model with the requirements of classical measurement theory, which places an emphasis on discriminant validity and prefers pure over confounded indicators with cross-loadings. Empirical research has made a big step forward during the last years, however, and this was also the reason why we could specify an acceptably fitting model with five values, which are similar to those in other published models. As a psychologist, Schwartz was much more concerned with the internal structure of his value system than Inglehart. Accordingly, it is not easy to derive hypotheses about the effects of values on attitudes and behavior from the theory of Schwartz. In this respect, the two-dimensional charts with the value circle can be misleading. Opposite values, for instance, do not always have opposite effects on external variables. Similar to suppressor variables, highly correlated adjacent values can completely change the signs of effects so that, as in our analyses, conservation and openness to change have the same effect on political orientations. Suppressor variables are usually discussed under the perspective that their omission has theoretically unpleasant effects. In our study, however, exactly the opposite is the case: The inclusion of adjacent values has turned the theoretically plausible negative correlation of openness to change and the left-right scale into a positive effect. A general lesson to be learned here is that small biases in the value measurement can dramatically change the effects of highly intercorrelated values on external variables and that a circular value structure is particular sensitive to this problem. This is one of the reasons why quite a few of these hypotheses were disconfirmed. So far the research on the relations with external variables seems to be in an exploratory stage.

Overall, both theories still are in a developmental stage. They both lack desirable methodological qualities. For the adherents of the theory of Inglehart it is a prior task to present reliable and valid measurement instruments. The traditional/secular-rational dimension in particular has not only to be more adequately operationalized but also deeper embedded in the theory. As far as the theory of Schwartz is concerned, the theory has to be elaborated especially in specifying more precisely the predictions concerning attitudes and behavior.

Acknowledgements

We are indebted to Eldad Davidov and two anonymous reviewers for helpful comments.
The work of Peter Schmidt was supported by the HSE Basic Research Program (International Laboratory for Sociocultural Research).
We thank Lisa Trierweiler for proof reading and language editing the manuscript.

References

Barnea, M. (2003) *Personal values and party orientations in different cultures*. Doctoral dissertation, The Hebrew University of Jerusalem, Israel.

Beckers, T., Siegers, P., & Kuntz, A. (2012) Congruence and performance of value concepts in social research. *Survey Research Methods* 6:13-24.

Beierlein, C., Davidov, E., Schmidt, P., Schwartz, S. H., & Rammstedt, B. (2012) Testing the discriminant validity of Schwartz Portrait Value Questionnaire items- A replication and extension of Knoppen and Saris. *Survey Research Methods* 6:25-36.

Bezonsky, M. D., Ciecuch, J., Durieux B., & Soenen, B. (2011) The how and what of identity formation: Associations between identity styles and value orientations. *Personality and Individual Differences* 50:295-299.

Bilsky, W. & Schwartz, S. H. (1994) Values and personality. European Journal of Personality 8, 163-181.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Caprara, G. V., Schwartz, S. H., Capanna, C., Vecchione, M., & Barbaranelli, C. (2006) Personality and politics: Values, traits, and political choice. *Political Psychology*27:1-28.

Costner H. L. & Leik R.K. (1964) Deductions from axiomatic theory. *American Sociological Review* 24:819-835.

Byrne, B. & van de Vijver, F.(2010) Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing* 10:107–132.

Davidov, E. (2010) Testing for comparability of human values across countries and time with the third round of the European Social Survey. *International Journal of Comparative Sociology* 51:171-191.

Davidov, E. (2008) A cross-country and cross-time comparison of the human values measurements with the second round of the European Social Survey. *Survey Research Methods* 2:33-46.

Davidov, E., Schmidt, P., & Schwartz, S. H. (2008) Bringing values back in: Testing the adequacy of the European Social Survey to measure values in 20 Countries. *Public Opinion Quarterly* 72, 420-445.

Davidov, E., Meuleman, B., Billiet, J., & Schmidt, P. (2008b) Values and support for immigration. A cross-country comparison. *European Sociological Review* 24:583-599.

Davidov, E. & Meuleman, B. (2012) Explaining attitudes towards immigration policies in European countries: The role of human values. *Journal of Ethnic and Migration Studies* 38:757-775.

De Graaf, N. D., Hagenaars, J., & Luijkx, R. (1989) Intragenerational stability of postmaterialism in Germany, the Netherlands and the United States. *European Sociological Review* 5:183-201.

Diamantopoulos, A., Riefler, P., & Roth, K. (2008) Advancing formative measurement models. *Journal of Business Research* 61:1203–1218.

Duriez, B., Van Hiel, A., & Kossowska, M (2005) Authoritarianism and Social Dominance in Western and Eastern Europe: The Importance of the Sociopolitical Context and of Political Interest and Involvement. *Political Psychology* 26:299-320.

Eagly, A. H. & Chaiken, S. (1993) *The Psychology of Attitudes*. Wadsworth, Fort Worth.

European Social Survey (2004) Round 2 Data, data file edition 3.2., 2011. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

Fishbein, M. & Ajzen, I. (2010) *Predicting and Changing Behavior: The Reasoned Action Approach*. New York: Psychology Press.

Flanagan, S. (1987) Changing values in advanced industrial societies revisited: Towards a resolution of the values debate. *American Political Science Review* 81:1303-1319.

Glen, N. D. (1974) Aging and conservatism. *Annuals of the American Academy of Political and Social Science* 415:176–186.

Hofstede, G. (2001) *Culture's Consequences. Comparing Values, Behaviors, Institutions, and Organizations across Nations*. Thousand Oaks: Sage.

Inglehart, R. (1977) *The Silent Revolution: Changing Values and Political Styles among Western Publics*. Princeton: Princeton University Press.

Inglehart, R. (1981) Post-materialism in an environment of insecurity. *The American Political Science Review* 75:880-900.

Inglehart, R. (1985) Aggregate stability and individual-level flux in mass belief systems: The level of analysis paradox. *The American Political Science Review* 79:97-116.

Inglehart, R. (1990) *Culture Shift in Advanced Industrial Society*. Princeton: Princeton University Press.

Inglehart, R. (1997) *Modernization and Postmodernization. Cultural, Economic, and Political Change in 43 Societies*. Princeton: Princeton University Press.

Inglehart, R. & Baker, W. E. (2000) Modernization, Cultural Change, and the Persistence of Traditional Values. *American Sociological Review* 65:19-51.

Inglehart, R. & Welzel, C. (2005) *Modernization, Cultural Change, and Democracy. The Human Development Sequence*. Cambridge: Cambridge University Press,.

Jagodzinski, W. (1983) Materialism in Japan reconsidered: Toward a synthesis of generational and life-cycle explanations. *American Political Science Review* 77:887-894.

Jagodzinski, W. (2004) Methodological problems of value research. In: Vinken, H., Soeter, J. and Ester, P. (eds.) *Comparing Cultures. Dimensions of Culture in a Comparative Perspective*. Leiden & Boston: Brill, 97-121.

Jagodzinski, W. & Manabe, K. (2009) On the similarity of religiosity in different cultures. In: Haller, M., Jowell, R. and Smith, T. W. (eds.), The International Social Survey Programme, 1984-2009. New York: Routledge, 313-336.

Jagodzinski, W. (2010a) Economic, social, and cultural determinants of life satisfaction: Are there differences between Asia and Europe?. *Social Indicators Research* 97:85-104.

Jagodzinski, W. (2010b) Is there a Conflict between Internal and External Validity Or: How Valid are Secular-Rational and Self-Expression Values? Paper presented at the Second European Value Studies Workshop after the 2008 Wave, University of Zagreb, June 7-11, 2010, retrieved from: http://info1.gesis.org/evs2008/Presentation [22.05.2012].

Jensen, J. P. & Bergin, A. E. (1988) Mental health values of professional therapists: A national interdisciplinary survey. *Professional Psychology: Research and Practice* 19:290-297.

Kluckhohn, C. (1951) Values and value-orientations in the theory of action: An exploration in definition and classification. In: Parsons, T. & Shils, E. (eds.) *Towards a General Theory of Action*. Cambridge (MA): Harvard University Press, 388-433.

Knoppen, D. & Saris, W. (2009a) Do we have to combine values in the Schwartz' Human Values Scale? A Comment on the Davidov studies. *Survey Research Methods* 3:91-103.

Knoppen, D. & Saris, W. (2009b) Schwartz's Theory of Human Values: Balancing Homogeneity of Reflective Items and Theoretical Coverage. RECSM Working Paper Number 9. Barcelona.

Marsh, H.W., Dowson, M.,Pietsch, J., & Walker, R. (2004) Why Multicollinearity matters: A Reexamination of Relations between Self-Efficacy, Self Concept and Achievement. *Journal of Educational Psychology* 3, 518-522.

Meredith, W. (1993) Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58:525-543.

Meuleman, B., Davidov, E., Schmidt P., & Billiet, J. (2013) Social location and value priorities. A European-wide comparison of the relation between socio-demographic variables and human values. In: Gabriel, O. and Keil, S. I. (eds.) *Society and democracy in Europe*. New York: Routledge, 45-67.

Norris, P. & Inglehart, R. (2004). *Sacred and Secular. Religion and Politics Worldwide*. Cambridge: Cambridge University Press.

Opp, K. D. and Wippler, R. (Eds.) (1990) Empirischer Theorienvergleich. Erklärungen sozialen Verhaltens in Problemsituationen. Westdeutscher Verlag, Opladen.

Piurko, Y., Schwartz, S. H., & Davidov, E. (2011) Basic personal values and the meaning of left-right political orientations in 20 countries. *Political Psychology* 32:537-561.

Rokeach, M. (1973*) The Nature of Human Values*. New York: The Free Press.

Sagiv, L. & Schwartz, S. H. (1995) Value priorities and readiness for out-group social contact. *Journal of Personality and Social Psychology* 69:437-448.

Sagiv, L. & Schwartz, S. H. (2000) Value priorities and subjective well-being: Direct relations and congruity effects. *European Journal of Social Psychology* 30:177-198.

Schmidt, P., Bamberg, S., Davidov, E., Herrmann J., & Schwartz, S. H. (2007) Die Messung von Werten mit dem 'Portraits Value Questionnaire' [The measurement of values with the 'Portraits Value Questionnaire']. *Zeitschrift für Sozialpsychologie* 38:261-275.

Schwartz, S.H. (1992) Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology* 25: 1-65.

Schwartz, S. H. (1994) Are there universal aspects in the content and structure of values? *Journal of Social Issues* 50:19-45.

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O. & Konty, M. (2012) Refining the theory of basic individual values. *Journal of Personality and Social Psychology* 103:663-688.

Schwartz, S. H. & Huismans, S. (1995) Value priorities and religiosity in four Western religions. *Social Psychology Quarterly* 58:88-107.

Schwartz, S. H. & Melech, G. (2000) National differences in micro and macro worry: Social, economic and cultural explanations. In: Diener, E. & Suh, E. M. (eds.), *Culture and Subjective Well-Being*. Cambridge (MA): MIT Press, 219-256.

Schwartz, S. H., Sagiv, L., & Boehnke, K. (2000) Worries and values. *Journal of Personality* 68:309-346.

Schwartz, S. H. & Rubel, T. (2005) Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology* 89:1010-1028.

Schwartz, S. H. (2006) Les valeurs de base de la personne: Théorie, mesures et applications [Basic human values: Theory, measurement, and applications]. *Revue Française de Sociologie* 47:929-968.

Schwartz, S. H. (2007) Value orientations: Measurement, antecedents and consequences across nations. In: Jowell, R., Roberts, C., Fitzgerald, R., & Eva, G. (eds.), *Measuring Attitudes Cross-Nationally. Lessons from the European Social Survey.* London: Sage, 169-203.

Schwartz, S. H. & Rubel-Lifschitz, T. (2009) Cross-national variation in the size of sex differences in values: Effects of gender equality. *Journal of Personality and Social Psychology* 97:171-185.

Schwartz, S. H. & Vecchione, M. (2011) A Revised Theory and New Instrument to Measure Basic Individual Values. Paper presented at the Regional Conference of the International Association of Cross-Cultural Psychology (IACCP), Istanbul, Turkey, June 30 – July 3, 2011.

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009) Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Quality & Quantity* 43:599-616.

Strupp, H. H. (1980) Humanism and psychotherapy: A personal statement of the therapist's essential values. *Psychotherapy* 17:396-400.

Veroff, J., Reuman, D., & Feld, S. (1984) Motives in American men and women across the adult life span. *Developmental Psychology* 20:1142-1158.

Verkasalo, M., Lönquist, J. E., Lipsanen J. & Helkama, K. (2009) European Norms and Equations for a two dimensional presentation of values as measured with Schwartz 21 item portrait values questionnaire. *European Journal of Social Psychology* 39:780-792.

Welzel, C. (2010) How selfish are self-expression values? A civicness test. *Journal of Cross-Cultural Psychology* 41:152-174.

Willis, G. B. (2005) *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks: Sage.

World Value Survey (2005) official data file v.20090901, 2009. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS, Madrid.

Appendix

Table A1

Measurement of the external variables, ESS 2004

Age	And in what year were you born?
Gender	Observation coding
Education	What is the highest level of education you have achieved? 1 = Primary or first stage of basic, 2 = lower secondary or second stage of basic, 3 = upper secondary, 4 = post secondary, non-tertiary, 5 = first stage of tertiary, 6 = second stage of tertiary
LEFT-RIGHT SCALE	In politics people sometimes talk of "left" and "right". Using this card, where would you place yourself on this scale, where 0 means the left and 10 means the right? (polintr)
POLITICAL INTEREST	How interested would you say you are in politics –are you… 1 = very interested, 4 = not at all interested? (polintr)
POLITICAL ACTIVISM	Additive index of 7 dichotomous items: Contacted politician or government official last 12 months (contplt), Worked in political party or action group last 12 months (wrkprty), Worked in another organisation or association last 12 months (wrkorg), Worn or displayed campaign badge/sticker last 12 months (badge), Signed petition last 12 months (sgnptit), Taken part in lawful public demonstration last 12 months (pbldmn), Boycotted certain products last 12 months (bctprd), 1=yes, 2=no
OPENNESS TO IMMIGRATION	CFA measurement model, 3 indicators: To what extent do you think [country] should allow people of the same race or ethnic group as most [country's] people to come and live here? (imsmetn), How about people of a different race or ethnic group from most [country] people? (imdfetn), How about people from the poorer countries outside Europe? (eimrcnt), 1 = allow many to come and live here – 4 = allow none
LIFE SATISFACTION	All things considered, how satisfied are you with your life as a whole nowadays? (stflife) 1=extremely unsatisfied, 10=extremely satisfied
CHURCH ATTENDANCE	Apart from special occasions such as weddings and funerals, about how often do you attend religious services nowadays? 1=every day, 7=never
ATTITUDES TOWARD GENDER EQUALITY	Men should have more right to a job than women when jobs are scarce (mnrgtjb) 1=agree strongly, 5=disagree strongly

Notes: The variable names in parentheses are those used in the original data set.

Table A2

Measurement of the external variables, WVS 2005

Age	Can you tell me your year of birth, please? (V236), This means you are ____ years old. (V237)
Gender	Observation coding (V235).
Education	What is the highest educational level that you have attained? (V138) 1 = no formal education, 2 = incomplete primary school, 3 = complete primary school, 4 = incomplete secondary school: technical/vocational type, 5 = complete secondary school: technical/vocational type, 6 = incomplete secondary: university-preparatory type, 7 = complete secondary: university-preparatory type, 8 = some university-level education, without degree, 9 = university-level education, with degree
LEFT-RIGHT SCALE	In political matters, people talk of "the left" and "the right." How would you place your views on this scale, generally speaking? (V114) 1=left, 10=right
POLITICAL INTEREST	How interested would you say you are in politics? Are you … 1 very interested, 2 somewhat interested, 3 not very interested, 4 Not at all interested (V95)
POLITICAL ACTIVISM	Additive index of 2 items: Joining in boycotts (V97), Attending peaceful demonstrations (V98), 1=have done, 2=might do, 3=would never do
OPENNESS TO IMMIGRATION	How about people from other countries coming here to work. Which one of the following do you think the government should do? (V124) 1 Let anyone come who wants to, 2 Let people come as long as there are jobs available, 3 Place strict limits on the number of foreigners who can come here, 4 Prohibit people coming here from other countries
LIFE SATISFACTION	All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are “completely dissatisfied” and 10 means you are “completely satisfied” where would you put your satisfaction with your life as a whole? (V22). 1 Completely dissatisfied, 10 Completely satisfied
CHURCH ATTENDANCE	Apart from weddings and funerals, about how often do you attend religious services these days? (V186.)1=more than once a week, 7=never, practically never
GENDER EQUALITY	When jobs are scarce, men should have more rights to a job than women (V44) 1=agree, 2=neither, 3=disagree

Notes: The variable names in parentheses are those used in the original data set.

Paper 3

The Concept of European Identity: Overused and Underspecified?

Co-authored with: Rossalina Latcheva, University of Zurich
Eva Rossbacher, University of Vienna

Published as: Latcheva, Rossalina, Datler, Georg, & Rossbacher, Eva (2012) The Concept of European Identity: Overused and Underspecified? In Salzborn, Samuel, Davidov, Eldad & Reinecke, Jost (eds.) Methods, theories, and empirical applications in the social sciences. Festschrift for Peter Schmidt. VS Verlag, 235-248.

The Concept of European Identity: Overused and Underspecified?

Introduction

Strengthening its citizens' identification with the process of European integration in general and with EU institutions and policies in particular has become one of the major challenges for the European Union today. The increasing diversification of European societies, a result of processes of globalisation and new kinds of differentiation and migration, forces political actors at different levels – local, regional, national and EU – to address issues of citizenship and identity as a means to counteract social conflict and the fragmentation of societies. Identity politics on the European level, i.e. attempts to create a collective identity through a redefinition of loyalties and attachments, aims at enhancing a consciousness of being European and thus at creating enduring forms of political and social integration.[1]

To gain information about the impact and extent of the identity building process with regard to the EU, the European Commission (EC) has been monitoring the evolution of public opinion in the Member States through surveys such as the Eurobarometer series, launched in 1973.[2] These studies are relevant to the preparation of texts, decision-making and the evaluation of European institutions' achievements. They inform policy makers, academics and media about people's expectations and concerns regarding Europe, their political attitudes, beliefs and behaviours (Wallström 2006). There is evidence of a broad consensus within the EC regarding the importance of opinion polling in the making of European democracy. Institutionalizing its integrative purpose in this sense promotes the emergence of a European political identity that could underpin trust in European institutions and policies and stimulate citizens' participation in the European political process. The importance of opinion polls and media monitoring has increased as has the tendency of citizens to withdraw from traditional politics (Commission of the European Communities 2006). *'Understanding European public opinion'* has therefore become one of the key targets as defined in the White

[1] Examples of contemporary European 'identity engineering' (Eriksen 2002) are for instance the euro-banknotes, the European flag, the anthem, 9th May as Europe Day and the motto of the EU *'United in diversity'*.

[2] For details see http://ec.europa.eu/public_opinion/index_en.htm

Paper on a European Communication Policy.[3] Recently, some alarming events (continuously decreasing acceptance of the EU in the member states, declining turnout at the elections to the European Parliament, the rejection of the draft for an EU constitution in France and the Netherlands) encouraged 'bottom-up' approaches like the "European Citizens' Consultations" in 2007 and 2009, which aim at *"closing the gap between the EU and its citizen by providing a platform for pan-European dialogue on the challenges facing the EU".*[4]

The concept of national and/or supra-national (i.e. collective) identity has also been the subject of many scholarly debates; it has been used both as an exogenous and as an endogenous dimension within different analytical frameworks. However, *the notion of European identity* as a concept may express several meanings. Some scholars frame *European identity* as a cosmopolitan orientation (Beck and Grande 2004, Delanty 2005), others advocate the development of a *"post-national constitutional patriotism"* (Habermas 1998), and still others talk about *"the cultural values of Europe"* (Joas & Wiegandt 2006). In particular, previous research has addressed the conceptualisation and determinants of individual national or supra-national identifications, and the interrelations between such individual identifications and forms of political trust, political participation and exclusionary reactions towards out-groups e.g. ethnic minorities and immigrants (McLaren 2006; Blank and Schmidt 2003; Coenders and Scheepers 2003).

Since identification with the European Union is seen as an essential prerequisite of sustaining the EU as a political entity and lending it greater popular support than it is enjoying at present and since *European identity* constitutes a core concept not only in political and scholarly debates but also in opinion polls and research programmes such as the Eurobarometer (EB), the European Value Survey (EVS) and the International Social Survey Programme (ISSP), it is of particular importance to regularly assess and adapt its conceptualizations and its measures. Notwithstanding the high-level of methodological consciousness within the above cited international data collection programmes, end-users of these data regularly insist on the need for more appropriate indicators for policy and comparative research. Yet, applying the concept of European identity presupposes sufficient information on what that identity consists

[3] The aim of the 2006 EC White Paper was the development of strategies for bridging the gap between the institutions of the EU and its citizens and the presentation of a broad range of techniques to better appraise European public opinion.

[4] For details see http://www.european-citizens-consultations.eu/

of, how it evolves over time and how different groups of Europeans share different aspects of European identity.

This research note aims to present a short synopsis of some recently developed insights about deficiencies in existing conceptualizations and measurements of *European identity*. In the first part we briefly discuss several theoretical approaches to European identity. In the second part we focus on current measurements of the concept and their potential flaws. Based on the analysis of 18 cognitive interviews conducted in Austria in 2003 and 2005 we also outline probable sources of measurement error with regard to the indicator of identification with Europe applied in the ISSP surveys 1995 and 2003. The paper ends with a short reflection on necessary prerequisites in the development of reliable and valid indicators of European identity for future research.

Conceptualizing *European Identity*

The idea that the capacity of the EU for 'identity building' is important to its effort to go beyond its technical/functional character and prove its strength as a social and political framework brings up the cognitive and normative dimensions of the 'identity building' process (Laffan 2004: 79). *"A normative lens on 'identity building' highlights the potential significance of shared values, norms and new roles in the European Union. A cognitive lens draws our attention to the importance of symbolic representations of the world"* (Laffan 2004: 78). At the same time, the growing "Europeanization of social relations" (Delanty and Rumford 2005: 54) has increased the possibilities for Europeans to construct and re-construct multiple local, regional, national and supra-national identifications. Following Tilly's approach of *relational realism* (2002), identities are social representations of social relations and can be defined as *"an actor's experience of a category, tie, role, network, group, or organization, coupled with a public representation of that experience; the public representation often takes the form of a shared story, a narrative"* (Tilly 2002: 80). Abdelal et al. (2006) specify four non-mutually-exclusive types of content regarding the meaning of collective identity and therefore distinguish not only between a normative and a cognitive dimension but also between a relational and a purposive content (cf. Table 1).

Table 1: Types of content of collective identities (Abdelal et al. 2006)

Constitutive norms (normative content)	Norms or rules that define group membership
Social purposes (purposive content)	Goals or purposes shared by an identity group
Relational comparisons (relational content)	Views and beliefs about other identities or groups, which are shared by an identity group
Cognitive models (cognitive content)	World views or understandings of political and material conditions and interests, which are shared by an identity group

The constitutive norms represent distinct practices associated with the process of identity building and the rules of membership in a specific collectivity. They can be formally encoded or informally enhanced. The content of a collective identity may be purposive when it is associated with certain collective goals, preferences and interests. Eriksen also describes identity formation on the EU level as conditional on individual or collective perceptions of utility: *"... enthusiasm for European integration is greatest in the poorest regions and among political elites, which may for different reasons reap the greatest economic and political benefits"* (2002: 75). The content of a collective identity is relational in the sense that it entails comparisons with and references to other collective identities, groups or historic periods. The cognitive content refers to world views, i.e. shared epistemological or ontological models. Following Abdelal et al. (2006), these contents are socially constructed and their meaning is always contested.

Some scholars suggest that analysis concerning individual identification with the EU should focus more on its social and political consequences, e.g. its effects on individual attitudes toward diversity and migration and not primarily on the definition of identity. Unfortunately, there is no consensus about the direction and the size of the effects of identification with the EU or with Europe on attitudes towards so-called *others* in the literature. Following some of the theoretical arguments, we might expect both a positive and a negative effect of identification with Europe on attitudes toward out-groups. On the one hand, scholars who frame European identity as cosmopolitan orientation (Beck & Grande 2004, Delanty 2005) would argue that identification with Europe is intrinsically linked with tolerance and liberal democratic attitudes. On the other hand, if we follow the line of reasoning within *Social Identity Theory* (Tajfel 1978), we would expect that processes of self-identification, self-categorization and social comparison may result in in-group bias, i.e. identification with Europe would lead to devaluation of individuals, groups, collectivities that are perceived as belonging to subjectively defined out-groups. Table 2 summarizes empirical studies that refer to *identification with Europe* as an explanatory variable for attitudes towards out-groups. The

results of these studies also point to differential direction of effects: some studies confirm a positive relationship between identification with Europe and tolerance towards different out-groups while others find a negative effect. Licata and Klein (2002) e.g. present evidence that identification with Europe decreases the tolerance towards foreign residents. The authors explain this finding with the specific subjective definition of *foreignness* in Belgium, in that foreign residents from non-European countries in Belgium are generally perceived as belonging to the out-group. In a series of experiments Mummendey, Waldzus and colleagues demonstrate that the negative effect of identification with Europe on attitudes towards out-groups can be explained by in-group projection. When respondents perceive their own nationality (e.g. being German) as a requisite for being European (the superordinate category), the social distance to so-called *others*, e.g. Poles, increases. This means, that individual identification with Europe leads to devaluation of out-groups when it is constructed from a national perspective.

Table 2: European Identity as a Variable. The relation with attitudes towards out-groups

Study	Design, Data	Independent Var.	Dependent Var.	Control / Manipulated Var.	Effect
Licata and Klein (2002)	Self-administered questionnaire, N=313, student sample, Belgium	Identification with Europe	tolerance towards resident foreigners	National identification, political orientation	-
Mummendey and Waldzus (2004) Waldzus et al. (2003) Waldzus and Mummendey (2004)	Several laboratory and web-based experiments, N=63, N=121, N=40	Dual identification: Europe and nation	Evaluation of the out-group (other European nationality)	Complexity of the representation of superordinate category, similarity of in-group and out-group, relative prototypicality of the in-group	- conditional on in-group projection
Citrin and Sides (2004)	Eurobarometer 53.0 (2000)	Identification with Europe	Attitudes towards minorities	Extensive multivariate control: national identification, age, gender, education, income, ideology, minority status etc.	+
Quintelier and Dejaeghere (2008)	Belgian Youth Survey 2006, 16 year old Belgians, N=4443	Feeling of belonging to the EU	Tolerance	Gender, education, contact with minorities,	+
Fuß (2006)	Regional samples of young adults (18-24) in Chemnitz and Bielefeld, Germany, N=2x400	European Identity	Attitudes to cultural/ national diversity	National identity	+

Delanty and Rumford (2005: 51) point to the complexity of the identity building process and emphasize four salient aspects of identity, which should be considered in any discussion of the concept and which also summarize both the arguments made until now and the most up-to-date findings with regard to identity: 1) identity arises only in relation to social action and is processual or constructed, 2) identities have a narrative dimension which pertains its

performative and public aspect, i.e. identity can be conceived as a discursive mode of self-understanding. Thus, the role of language becomes significant for the shaping of identities; 3) identities are constructed at the interface between self-images and images of the other and in this sense they are based on the creation of difference and commonality; they exist in a relational context which under the conditions of modernity entails reflexivity; 4) identities can be overlapping (as in hyphenated identities), nested, cross-cutting, mixed (or hybrid) or co-existent. Following the latter, identifications with the European Union are not necessarily incompatible with local, regional, national or ethnic identities (cf. also Eriksen 2002: 76).

Yet, European identity has been primarily conceptualised and measured in opposition to national identity or solely as attachment to the EU, and the existence of multiple attachments and their gradual, relational and discursive (performative) character has not yet been sufficiently taken into account. As a result, deficiencies in conceptualizing European identity will particularly influence the quality of its measurement and consequently the ability to draw conclusions about social reality. EB and other relevant surveys such as the European Social Survey (ESS), ISSP, EVS etc. might make more substantial contributions to the knowledge base of European policy makers, if the validity and cross-cultural equivalence of the measurements of European identity are thoroughly developed and continually optimized.

Measuring European identity and its deficiencies

The theoretical and empirical approaches discussed so far exemplify that *"the notion of identity means quite different things to different people, and even when the concept is commonly defined, measuring it remains extremely difficult"* (Hermann and Brewer 2004: 4). Despite the fact that there has been a constant effort to improve the quality and significance of European identity indicators, there is considerable evidence suggesting a theory-driven reassessment of the methodology that has been used so far. What is missing is a theoretical and methodological update for a variety of indicators such as emerging multiple identities, EU scepticism and public preferences concerning policy ends. Table A1 in the appendix presents the operationalizations of identification with Europe/EU in the studies discussed above. Although most of them use multiple indicators they refer to *identification with Europe* as a one-dimensional concept. The differences between the applied measures are rather a matter of degree than of conceptual dissimilarity. Many items refer to "feeling", "feeling attached" or "being proud of", i.e. to emotional attachment. Still others, including the one used in Eurobarometer until 2008, ask about "being" or "seeing oneself", i.e. about self-conception or self-categorization as European. A partial exception is the *European Citizenship Scale* by

Quintelier and Dejaeghere (2008), which covers the political dimension of the identification with Europe. Other authors occasionally use the term *European citizen* in their operationalizations.

Nevertheless, in most of the studies the terms *Europe*, *EU* and *European* are used without explicitly specified meanings. Given the contested nature of what European identity ought to be, this is highly problematic. The relation between identification with Europe/EU and exclusionist attitudes will largely depend on whether respondents frame it as cosmopolitan and democratic or whether Europe is seen as a fortress against immigration. It must however been born in mind that opting for a liberal and democratic Europe may still go hand in hand with an endorsement of Europe as a fortress (in the sense that the benefits of the Union should stay inside its borders).

In the following we exemplify some of the problems that might arise, at least in the Austrian context, if one applies the 1995/2003 ISSP item (cf. Table A1) as a measurement of identification with Europe. Our analysis is based on qualitative material from 18 cognitive interviews with Austrian citizens, carried out between 2003 and 2005 (6 in 2003 and 12 in 2005).[5] Since the early 1980s, cognitive interviewing has become indispensable in the development and testing of survey questionnaires (Willis 2005, Tourangeau 1984). It must be said, however, that a reformulation of items based on cognitive pre-tests does not necessarily lead to indicators that are more valid (see e.g. Schüßler and Schmidt 2008). Reformulated items have to be re-tested as well. In an earlier multi-method study, which aimed at assessing the measurement quality of the ISSP items measuring *national pride* and which used the same qualitative material, Latcheva (2009) has shown that respondents adopt different perspectives while answering the same questions and that this generates incomparability of answers to the extent that respondents with diametrically opposed attitudes opt for the same response category. Moreover, it appears that different respondents attribute different meanings to key terms (e.g. *democracy*, *history*), which in turn results in a significant amount of measurement error or in missing data.

The focus in the analysis in this paper is on the subjective definitions of the term *Europe* and how respondents construe their European identification while answering the ISSP question:

[5] Respondents were selected along the following criteria: gender, age (<=30; 31-55; =>56), education (with/without university entrance qualification) and urban vs. rural context. For more details on design and an example of an integrated analysis of cognitive interviews and factor analytic techniques, see also Latcheva 2009.

"How close do you feel to Europe?" and the probes that follow the original question. For this specific item, we applied two cognitive techniques: *category selection probing* for the original ISSP question (e.g. "Why did you select this answer category?") and *special comprehension probing* (e.g. "What does *Europe* mean to you?"). A very brief summary of the varying perspectives that our respondents adopted when answering these questions is presented in Table 3.

Table 3: Perspectives respondents adopted when answering the questions

How close do you feel to Europe	**very close to close feeling due to:** perceptions of common mentality; common ideas; a common social policy; the community per se; the diversity of Europe; the nearness of different countries, freedom of mobility **not so close or not close at all feeling due to**: "What is meant by Europe anyway?"; Europe is just an idea, an abstract construction but not a fact; Europe as a political formation; Europe is too global and for that reason not able to inspire any emotional attachment;
Subjective definitions of the term 'Europe'	**Europe defined as:** a continent; Europe=EU; Europe as a *Lebenswelt;* a beautiful land; home; an entity with a common historical and religious background*;* just a few, hardly connected countries; just a vague idea

Positive feelings towards Europe are based both on what Europe represents as an overall entity and on its diversity: on the one hand, the respondents name the common ideas, a common social policy and the community as a whole (so referring to Europe in the sense of the EU); on the other hand, the answers are related to the diversity of languages, peoples, cuisines, landscapes. Not so close or not at all close feelings towards Europe derive from the attitude that Europe is seen just as vague idea or as a more or less artificial or constructed, not a historically grown entity. The decision in favour of a certain answer category often depends on comparison with other countries, continents, groups or points in time. This way of reasoning corresponds to the relational content of a collective identity outlined by Abdelal et al. (2006, cf. Table 1) and to Tajfel's (1978) definition of social identity since most of the respondents emphasize the positive aspects of specific groups, countries or continents they feel close to while slightly devaluating all those categorized as *others.* Two specific findings are of particular importance with regard to the ISSP item *"How close do you feel to Europe?"*: 1) the definition of the term *Europe* is too broad and unspecific – this produces variance in the question's interpretation (i.e. not a valid variance) and also item-non-response (i.e. selection of "can't choose" as an answer); 2) the term "close to" introduces an emotional note that some of our respondents felt is inappropriate with regard to Europe or the EU. This diminishes the quality of the answers and the respondents' motivation to conduct the necessary cognitive work.

We observed also numerous subjective definitions of Europe, which conceive of Europe as a continent, as identical to the EU, as *Lebenswelt* and an entity, but also as a vague idea or skeletal formation, and which are strongly suggestive of the answers to the question of how close one feels to Europe. Although some of the respondents explicitly decided to see themselves either as European or as Austrian, the measurement of European identification in contrast to national identification, i.e. using a dichotomy instead of allowing multiplicity and degree of individual attachments, seems at least inappropriate.

Conclusions

The 18 cognitive protocols suggest that identification with Europe is, as opposed to that with a nation, based on rationale, utility and content rather than on emotion. Thus, being a member of a specific ethnic or religious group does not preclude being Austrian; and being Austrian does not preclude being European etc. Hermann and Brewer (2004) argue that the difficulty to measure identity arises from the fact that people belong to several communities simultaneously and the relative significance of these attachments can change over time and space and according to social context and the issues or the contents addressed in it. Since most indicators within the surveys cited in this paper consist of deduced measurements of the underlying theoretical concept *European identity* and since a variety of significant theoretical differentiations have been introduced in recent years, it seems appropriate to make this progress fruitful for a continuous improvement of the standard indicators and scales that have been used to measure European identity so far.

According to our results we can summarize that new indicators which depict multiple contents and relational contexts in the measurement of individual identification with Europe and which build on e.g. public preferences with regard to policy ends, membership, benefits, merits, values and norms, but not on distilled emotional components (as in the measurement of national identification), are to be strongly recommended (cf. also Schmitt 2003: 248). Since such a set of indicators is more likely to cover the social representations of Europe or of *being European*, they might yield new, high-quality evidence on the development of European identity across the EU and its significance for strengthening individuals' trust in European institutions and the policy making process on the EU level.

References

Abdelal, R., Herrera, Y. M., Johnston, A. I. & McDermott, R. (2006) Identity as a Variable. *Perspectives on Politics* 4:695-711.

Beck, U. & Grande, E. (2004) *Das kosmopolitische Europa. Gesellschaft und Politik in der zweiten Moderne.* Frankfurt am Main: Suhrkamp.

Blank, T. & Schmidt, P.(2003) National Identity in a United Germany: Nationalism or Patriotism? An Empirical Test With Representative Data. *Political Psychology* 24:289-312.

Citrin, J. & Sides, J.(2004) More than Nationals: How Identity Choice Matters in the New Europe. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 161-185.

Coenders, M., Lubbers, M. & Scheepers, P. (2003) Majority populations' attitudes towards migrants and minorities. Vienna: European Monitoring Centre on Racism and Xenophobia (EUMC), Ref. no. 2003/04/01.

Commission of the European Communities (2006) White Paper on a European Communication Policy. Brussels: COM (2006) 35 final.

Delanty, G. & Rumford. C. (2005) *Rethinking Europe. Social theory and the implications of Europeanization.* London: Routledge.

Delanty, G. (2005) The Idea of a Cosmopolitan Europe. *International Review of Sociology* 15: 405-421.

Dejaeghere, Y. & Quintelier, E. (2008) Does European Citizenship Increase Tolerance in Young People? *European Union Politics* 9:339–362.

Eriksen, T. H (2002) *Ethnicity and Nationalism: Anthropological Perspectives*. London: Pluto Press.

Fuß, D. (2006) *Europa als Quelle sozialer Identität*. Dissertation, University of Gießen, Germany.

Habermas, J. (1998) *Die postnationale Konstellation. Politische Essays*. Frankfurt am Main: Suhrkamp.

Hermann, R. K. & Brewer. M. B. (2004) Why Study Identity and Institutions? In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 247-271.

Joas, H. & Wiegart, K. (eds.) (2006) Die kulturellen Werte Europas. Frankfurt a. M.: Fischer.

Laffan, B. (2004) The European Union and Its Institutions as 'Identity Builders'. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 75–96.

Latcheva, R. 2009. Cognitive Interviewing and Factor-Analytic techniques: A Mixed Method Approach to Validity of Survey Questions Measuring National Identity. *Quality & Quantity* 45:1175-1199.

Licata, L., & Klein, O. (2002) Does European citizenship breed xenophobia? European identification as a predictor of intolerance towards immigrants. *Journal of Community and Applied Social Psychology* 12:1-15.

Mclaren, L. M. (2006) *Identity, Interests and Attitudes to European Integration*. Basingstoke: Palgrave.

Mummendey, A. & Waldzus, S. (2004) National Difference and European Plurality: Discrimination or Tolerance between European Countries. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 59-72.

Schmitt, H. (2003) The Eurobarometers: Their Evolution, Obvious Merits, and Ways to Add Value to them. *European Union Politics* 4:243-251.

Schüßler, A. & Schmidt, P. (2008) Does cognitive interviewing matter? Testing the effects of cognitive methods on data quality. Paper presented at the 7th International Conference on Social Sciences Methodology (RC 33, ISA), Neapel, Italy.

Taifel, H. (1978) *Differentiation between Social Groups*. London: Academic Press.

Tilly, C. (2002) *Stories, Identities, and Political Change*. Lanham: Rowman & Littlefield.

Tourangeau, R. (1984) Cognitive science and survey methods. In: Jabine, T. B., Straf, M. L., Tanur, J. M. & Tourangeau, R. (eds.) *Cognitive Aspects of Survey Methodology. Building a Bridge Between Disciplines*. Washington: National Academic Press, 73-100.

Waldzus, S., Mummendey, A., Wenzel, M. & Weber, U. (2003) Towards tolerance: Representations of superordinate categories and perceived ingroup prototypicality. *Journal of Experimental Social Psychology* 39:31–47.

Waldzus, S. & Mummendey, A. (2004) Inclusion in a superordinate category, in-group prototypicality, and attitudes towards out-groups. *Journal of Experimental Social Psychology* 40: 466–477.

Wallström, M. (2006) Understanding public opinion. Madrid: Opening speech at the stakeholders' conference on "Understanding public opinion", 27 October 2006.

Willis, G. B. 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. London: Sage.

Appendix

Table A1: Measures of identification with Europe

Licata and Klein (2002): 9 item scale 1. "I feel attached to the European ground." 2. "I would miss Europe if I had to leave it." 3. "My destiny is linked with that of other Europeans." 4. "I feel solidarity with all the other Europeans." 5. "I am proud to tell my friend that I am European." 6. "There is no reason to be proud to be European." 7. "Being European is something I rarely think about." 8. "I don't have clear feelings about the fact that I am European." 9. "Most of the time I like to think about myself as a European." [1 "totally disagree" - 11 "totally agree"]
Mummendey and Waldzus (2004), Waldzus et al. (2003), Waldzus and Mummendey (2004): 6 item scale 1. "I identify with Europeans.'' 2. ''I have a negative attitude towards Europe.'' 3. ''I consider myself to belong to Europeans.'' 4. ''I like being European.'' 5. ''I feel like a European.'' 6. ''I feel sceptical about a European identity".' [1 "not at all" - 5 "very much"]
Quintelier and Dejaeghere (2008): 4 item ''European Citizenship Scale" 1. "How much trust do you have in the European Parliament?" 2. "How far-reaching or important are the decisions of the European Union for your daily life?" 3. "I consider myself in the first place a European citizen." 4. "Thanks to the European Union I feel myself a full citizen of Europe." [Response scale: not documented in Quintelier and Dejaeghere (2008)]
Fuß (2006): 3 items, Confirmatory Factor Analysis measurement model 1. "Now I would like to ask about the strength of how you feel about being different sort of nationality? On a scale of 0-4 how do you feel about being ... "European" [0 "no feeling at all" - 4 "very strong feeling" 2. "How would you rate the importance of the following in terms of who you are, that is, how you feel or think about yourself as a person?" ... "being a citizen of the European Union" [0 "not at all important - 4 "very important"] 3. "People may feel different degrees of attachment to their city, town or village, to their region, to their country or to Europe. Thinking about your own attachment, and using the scale on this card, please tell me how attached you feel, to" ... "Europe" [0 "not at all attached" - 4 "completely attached"]
Eurobarometer (EB) **EB 53.0 (2000),** used by **Citrin and Sides (2004)** In the near future, do you see yourself as …? [NATIONALITY] only / [NATIONALITY] and European / European and [NATIONALITY] / European only **New item introduced in EB 69.2 (2008)** Thinking about this, to what extent do you personally feel you are European? ["to a great extent" – "somewhat" – "not really"– "not at all"]
International Social Survey Programme (ISSP) 1995/2003, Module National Identity I & II How close do you feel to … "Europe" ["very close" – "close" – "not very close" – "not close at all"]
European Value Study (EVS) 1999/2000 Which of these geographical groups would you say you belong to first of all? And the next? And which do you belong to least of all? C Your country as a whole D Europe A Locality of town where you live B Region of county where you live E The world as a whole

Paper 4

European identity as a safeguard against xenophobia? A differentiated view based on identity content

---

Forthcoming in: Dynamics of National Identity. edited by Grimm, Jürgen, Huddy, Leonie, Schmidt, Peter and Seethaler, Josef. London: Routledge.

European identity as a safeguard against xenophobia? A differentiated view based on identity content

Abstract

Prominent theorists hope that European identity could do without the drawbacks of national identity, not the least of which is hostility towards foreigners. However, empirical evidence of whether or not European identity is a safeguard against xenophobia is inconclusive. Whereas the theoretical claim rests on a specific substantive conception of European identity, i.e., on a specific content of European identity, the measures commonly used refer to identification with Europe only. Such measures do not tap the content of European identity at all. The paper employs a novel strategy to bring identity content into quantitative models: a latent class factor model. On the basis of 2009 Eurobarometer data, three dimensions of identity content ('social democracy', 'openness & freedom', and 'the negative') are distinguished. The relation of identification with Europe and attitudes towards foreigners varies substantially once identity content is taken into account. It therefore depends whether or not European identity is a safeguard against xenophobia: it depends on the content of European identity.

European identity is a prominent concept. A brief glance at recent book titles suggests a stable bull market (e.g., Bruter, 2005; Checkel and Katzenstein, 2009; Cerutti and Lucarelli, 2008; Fligstein, 2008). European identity is an attractive concept; scholars display their sympathy for European integration. National identity is a prominent concept as well, but a much less attractive one.

This paper first clarifies the conceptual status of European identity, starting from a comparison with national identity. In particular, I sketch which consequences collective identities have for non-members (*their side effects*), describe why political theory nevertheless acknowledges collective identities (*the supposed ends*), and introduce the claim that European identity could do without the drawbacks of national identity (*the hopes*). In a second step, I outline an analytical framework for the empirical study of European identity that emphasizes identity content. The third step reconstructs the argument why European identity could be a safeguard against xenophobia. Previous research on this issue has yielded contradictory

results, which is—as I will argue—probably due to an inappropriate conception and measurement of European identity. I propose a novel strategy to capture different types of identity content and apply it to Eurobarometer data: a latent class factor model. The resulting classification represents different types of European identity content, which allows studying the relationship of European identity and attitudes towards foreigners as being conditional on the meaning of European identity.

European and national identity

European identity and national identity have many things in common. Both are at once concepts of practice and concepts of analysis (Brubaker and Cooper, 2000). In politics, they are used to support claims about how societies should be. In science, they are used to study how societies are and how they alter. This dual use creates tension—and this is where national identity and European identity separate. Whereas scholars of national identity distance themselves from nationalism, many scholars who investigate European identity simultaneously express their sympathy for European integration (e.g., Delanty and Rumford, 2005; Habermas, 1998).

This distinction can be understood in terms of *the side effects* and *the supposed ends* of all collective identities and *the hopes* that are associated exclusively with a European identity specifically. The danger of side effects is inherent in collective identities. The construction of in-groups is the flip side of the construction of out-groups— to identity 'us', 'we' need 'them' (Jenkins, 1996). Ever since the invention of the nation, non-members have suffered (see, e.g., Hobsbawm, 1992; Wimmer, 2002). Nevertheless, political theorists are reluctant to expel collective identity from their toolkit because they regard it as a prerequisite for democracy. If there were no collective identity, there would be no demos, no people. If there were no people, democracy—the rule of the people—would not be possible (see, e.g., Weiler, 1999). With nationalism discredited by history, the hope is that Europeanism could serve the same ends as nationalism, albeit without the latter's side effects. When scholars speak of a cosmopolitan turn (Beck and Grande, 2004) or post-national patriotism (Habermas, 1998), they hope that a European identity will not involve the denigration of others but tolerance and respect.

European identity is a screen onto which many project their hopes. Although it may be attractive as an object of normative projections, its vagueness renders it not very useful for analytical purposes. The next section develops a more qualified conception of European

identity. On this basis, we can then reconstruct the argument why European identity might be able to do without the side effects of national identity.

A conceptual framework for the empirical study of European identity

'Identity' is a controversial concept; some want to get rid of it (Brubaker and Cooper, 2000), others want to get it right (Tilly, 2003). At least it is a demanding and tricky concept. Let us start—without the prefix 'European'—from a simple vantage point: Identity is the answer to the question who we are. Consider two exemplary answers:

(1) 'I am an aquarium hobbyist.'

(2) 'I am Austrian.'

The first answer stresses a characteristic of mine; its focus is on the alleged difference between me and you. The second answer stresses my membership in an imagined group; its focus is on the alleged similarity of me and my fellow Austrians. Roughly along this line, it is common to distinguish 'personal' identity, which stresses distinctiveness, from 'social' or 'collective' identity, which stresses sameness (for an overview, see Fearon, 1999). If taken literally, splitting the personal from the social is problematic: on the side of personal identity there is a danger to think of a person as highly individualized and unaffected by social relations ('undersocialized'); on the other side lurk collectives endowed with feelings and memories that govern the individual (anthropomorphism, 'oversocialized'). It is more appropriate to acknowledge that it is always the individual who experiences identity. A collective neither thinks nor feels. Rather, the identity of every individual is made up of aspects that make her different and of aspects that he shares with others. Similarity and difference rest on the comparison with others and are sensitive to context— imagine an aquarist club in Zurich, in which case statements (1) and (2) would reverse functions. Identity is always social identity (Jenkins, 1996).

Brubaker and Cooper (2000) devise a new inventory of terms to grasp that identity has individual aspects (*identification*) and collective aspects (*commonality, connectedness, and groupness)*.[1] Many researchers focus exclusively on identification, for it seems conceptually

[1] Brubaker and Cooper (2000) propose other terms as well: *self-understanding* and *social location* (as dispositional terms, to explain non-instrumental action). Another important aspect of identity is *external categorization*, i.e., others who identify us. This potentially coercive aspect is considered less relevant to European identity (for the dialectics of internal and external identification, see Jenkins 1996).

clearer than the collective facets (see, e.g., Opp, 2005).[2] However, a closer look at the process of identification will illustrate that we need to take into account the collective aspect of identity to make sense of individual identification.

The process of identification has three components. First, it requires a *subject* who identifies with an *object* (Gerhards, 2000). This is straightforward. But consider the question why a person identifies with a group or a social category. We cannot give an answer unless we consider what the object means to the subject who identifies with it. In this vein, Abdelal et al. (2009) propose to describe identities in terms of *content and contestation*. Content refers to the shared, collective aspect of identity and includes the norms, goals, and beliefs that make up the meaning of an identity. What meaning an identity has is neither fixed nor given but contested. Therefore, contestation highlights that individuals can disagree over the content of an identity. Both content and the degree of contestation invite empirical investigation. Analytically, we can separate identification from the collective aspects of identity, but we cannot grasp why people identify with an object when we do not know what this object means to them.

What applies to identity in general also applies to European identity. European identity is used to denote different phenomena at the collective level (e.g., ideas of the European polity) and at the individual level (e.g., feelings of attachment to Europe). The individual and the collective level—i.e., identification with Europe and the content of European identity—are interrelated (e.g., 'I feel European because I share a certain conception of the European polity', or 'As I feel European, I'm willing to accept a European political community').

But when people identify with Europe, what is the object that they identify with? 'Europe' is a polyvalent term that denotes geographical regions, refers to history, culture, and to the political project of the EU, amongst other things. Actually, there are several potential objects for the identification with Europe. This is how complex the social reality of the supposed-to-be-one-thing 'European identity' is. However, one object stands out in several respects: it is much less questionable what the EU is as compared to 'European culture' or 'European homeland.' After all, it remains an empirical question whether many people give the EU a prominent place in their conception of Europe.

[2] Another reason may be that identification fits well with methodological individualism and attitudinal survey research.

Is European identity a safeguard against xenophobia?

When theorists such as Habermas and Derrida (2004) talk about European identity, they are talking about a European identity with a specific content. Though they offer good reasons why this content should be post-national patriotism or trans-national solidarity and explain why the socio-cultural history of Europe makes it ripe for developing such an identity, they neither investigate nor claim what is the case. Politically, this means they engage in constructing the content of a European identity. Analytically, this means their hope that a European identity would safeguard against xenophobia is conditional upon the content of European identity.

Surprisingly, identity content is largely absent in studies on European identity and xenophobia. Most studies make reference to social identity theory (SIT) (Tajfel and Turner, 1986). However, these studies only employ a very basic notion of SIT—critics have called it 'SIT-lite' (McGarty 2001) or even a misreading of SIT (Reicher 2004). The simplified SIT story goes like this: Group membership leads to identification with the ingroup, which in turn leads to the devaluation of relevant out-groups. The main reason for this biased comparison is members of the ingroup striving for a positive self-conception. Based on SIT-lite, two arguments for the tolerance-enhancing effect of European identity are put forward:

(1) Group size: European identity is more encompassing than national identity. Foreigners are part of the ingroup (Fuß, 2006; Gaertner et al., 1993; Opp, 2003; Skrobanek, 2004).
(2) 'Contagion': The European ingroup contains EU foreigners. The positive effect is 'contagious' to non-EU foreigners (Dejaeghere and Quintelier, 2008).

The argument of group size is imprecise; it is not clear whether the effect is only meant to hold for EU foreigners or for all foreigners. Contagion posits that the inclusion of EU foreigners in the ingroup would transmit to all foreigners. Essentially, both arguments rely on the argument of size—Europe is bigger than the nation.

SIT was developed in lab experiments (e.g., participants saw paintings and were randomly (!) assigned to a Klee group or a Kandinsky group). Under these conditions—which became famous as the 'minimal-group paradigm'—group size may be the main factor for intergroup processes. How does that translate to the real world?

Reicher (2004) argues that the minimal-group experiments could equally well be described as 'maximum-group experiments'—for the lab condition creates maximally defined social identities: Lab identities are arbitrarily created, unambiguous categories, with clear group boundaries and fixed identity content. Political identities in the real world, however, are

different—they are ambiguous, meaningful, and contested (Huddy, 2001). Therefore, their content needs to be considered, especially when it comes to understanding the consequences of group identification (Billig, 1995; Reicher and Hopkins, 2001: 77-99).

Studies on the relation of European identity and attitudes towards outgroups have yielded contradictory results. Table 1 gives an overview.

Table1: The relation of European identity and attitudes towards outgroups

Study	Design, data	Independent var.	Dependent var.	Control / manipulated var.	Effect	Effect size: stand. regr. coeff.
Citrin and Sides (2004)	Eurobarometer 53.0	Identification with Europe	Attitudes towards minorities	Extensive multivariate control: national identification, education, ideology, etc.	+	.08
Dejaeghere and Quintelier (2008)	Belgian Youth Survey, *N*=4443	Feeling of belonging to the EU	Tolerance	Gender, education, contact with minorities	+	.08
Fuß (2006)	Regional samples of young adults, Germany, *N*=2x400	European identity	Attitudes toward cultural/ national diversity	National identity	+	.11
Heyder and Schmidt (2005)	GFM-Survey, Germany, , *N*=3000	Identification with Europe	Attitudes towards foreigners	National and regional identification	+	.05*)
Licata and Klein (2002)	Student sample, Belgium, *N*=313	Identification with Europe	Tolerance towards resident foreigners	National identification, political orientation	-	.29
Mühler and Opp (2004)	Regional samples, Germany, *N* = 2x ≈ 1500	Identification with Europe	Attitudes towards foreigners	Extensive multivariate control: regional/national identity, perceived discrimination, political orientation, etc.	+	.13
Skrobanek (2004)	Regional sample of young adults, Germany, *N*=1457	European identity	Attitudes towards foreigners	Extensive multivariate control: regional/national identity, education, values, etc.	+	.13 – .15
Waldzus et al. (2003), Mummendey and Waldzus (2004)	Laboratory and web-based experiments, *N*=63, *N*=121, *N*=40	Dual identification: Europe and nation	Evaluation of the outgroup (other European nationality)	Complexity of the representation of superordinate category, similarity of ingroup and outgroup, relative prototypicality of ingroup	+ direct – indirect (conditional on ingroup projection)	+.11 –.20 –.07 – -.16*)

*) Indirect effect computed from the path coefficients in the publication. The model posits a negative effect of identification with Europe on hostility towards foreigners, mediated by patriotism. For reasons of consistency table 1 displays the equivalent positive effect on attitudes towards foreigners in table 1.

**) Indirect effects computed from the path coefficients in the publications. The total effect (direct effect + indirect effect) is small and positive, as in most other studies. Note that these studies, in contrast to all other studies, are experiments that manipulate conditions.

Of course different results can be—and in part will be—due to different methods. I suggest that another part of the differences can be explained by substantive reasoning. We have three

facts: (1) ‘Large-scale’ surveys always find positive correlations. (2) These correlations are weak. (3) Some ‘smaller’ studies, inconveniently, yield opposite results. Now imagine that European identity were not the same thing for different people. Imagine further that different European identities would have different relations with attitudes towards outgroups. In larger datasets with more representative samples, the results would mirror an average effect. Typically, average effects are low. Smaller studies might capture a specific European identity—either by chance or by design. The studies of Waldzus, Mummendey, and colleagues are key in this respect: whereas most of the affirmative studies survey European identity, their experimental design manipulates the characteristics of European identity. They demonstrate that, in case of ingroup projection, the relation of identification with Europe and attitudes towards outgroups turns negative. Ingroup projection describes a process by which individuals project the qualities of their ingroup onto the superordinate category. When individuals perceive ‘being German’ (ingroup) as prototypical for ‘being European’ (superordinate category), the social distance to ‘being Polish’ (outgroup) is increased precisely through the identification with Europe. Put differently, ingroup projection alters the content of European identity. Thus, I hypothesize that the relation between identification with Europe and attitudes towards foreigners depends on the content of European identity.

The problem: no content in the measures

Although some of the studies name their variables ‘European identity,’ these studies across the board use similar measures: measures of identification with Europe. All studies treat their concepts as uni-dimensional. Table 2 provides typical examples.

One type of item, including the one used as a standard measure in the EB until 2008, just asks about ‘being’ or ‘seeing oneself.’ Items of this kind refer to self-perception. Another type of item uses ‘feeling’ or ‘feeling attached.’ Such items refer to emotional attachment. The object is denoted by ‘Europe’, but the meaning of ‘Europe’ is not made explicit. The items measure the degree of identification with Europe without grasping the content of this identity. If the content of European identity is what determines its relation to attitudes towards foreigners, this is problematic.

Table 2: Measures of identification with Europe

Eurobarometer (EB) 53.0, used by **Citrin and Sides (2004)** In the near future, do you see yourself as …? [NATIONALITY] only / [NATIONALITY] and European / European and [NATIONALITY] / European only
Skrobanek (2004), Müller and Opp (2004) To what extent do you feel European?
Fuß (2006): 3 item-factor model 1. Now I would like to ask about the strength of how you feel about being different sorts of nationality? How do you feel about being ... European 2. How would you rate the importance of the following in terms of who you are, that is, how you feel or think about yourself as a person?' ... being a citizen of the European Union' 3. People may feel different degrees of attachment to their city, town, or village, to their region, to their country or to Europe. Thinking about your own attachment, please tell me how attached you feel, to ... Europe
Licata and Klein (2002): 9-item scale 1. I feel attached to the European ground. 2. I would miss Europe if I had to leave it. 3. My destiny is linked with that of other Europeans. 4. I feel solidarity with all the other Europeans. 5. I am proud to tell my friends that I am European. 6. There is no reason to be proud to be European. 7. Being European is something I rarely think about. 8. I don't have clear feelings about the fact that I am European. 9. Most of the time I like to think about myself as a European.

Addressing the problem: bringing content in

Bruter (2004, 2005) has suggested distinguishing two types of European identity: civic and cultural. Civic European identity refers to the identification with the political structure of the EU. Cultural European identity refers to the identification with an imagined group of Europeans, its history, values, or way of life. Following a similar approach, Pichler (2008) distinguishes a political from a cultural dimension.

While I do think that differentiating between different content of European identity is the right track, I doubt that this dichotomy of political and cultural will do the trick. It mirrors the classical distinction between civic and ethnic conceptions of the nation (Brubaker, 1994) and between patriotism and nationalism (Kosterman and Feshbach, 1989; Schatz et al., 1999). While the distinction of a more inclusive and liberal patriotism and a more exclusive and hostile nationalism have proven useful in some research contexts (Blank and Schmidt, 2003), the validity of this dichotomy is not without its critics, who argue that nationalism and patriotism are too similar or too interrelated to separate them (Bauman, 2000; Heinrich, this

volume). The underlying civic–ethnic distinction is now regarded as too simplistic and normative itself (Brubaker, 1999).[3, 4]

Instead of imposing the theory of nationalism on European identity, I propose an exploratory, empirical approach. I will not limit the types of European identity content to any specific number a priori and then search for measures of these dimensions; instead I will start from the data and model the minimum number of dimensions necessary to represent the data. Given the diversity of pictures of Europeanness, I predict the degree of contestation over the content of European identity to be high.

Data and measures

The various Eurobarometer (EB) studies frequently include questions on the identification with Europe, on the meaning of the EU, and on attitudes towards foreigners. In 2009, EB 71.3 (European Commission 2012) addressed these three topics in a body. I used the EU-citizen samples of the by-then EU27 member states, N=26,830.[5] The proportion of missings in the measures is small (<2%); therefore listwise deletion is acceptable (Graham, 2009). Proportions and other descriptive statistics were weighted according to the country's population size. Table 3 shows the measures.

[3] Yack (1998: 203), paraphrasing Hobbes, coined the aphorism 'Nationalism is patriotism misliked and patriotism, nationalism liked' in a polemic against Maurizio Viroli, the author of *For Love of Country: An Essay on Patriotism and Nationalism.*

[4] Roccas and Berlin (this volume) provide an overview of various models of the dimensional structure of identification.

[5] Information on sampling and fieldwork are available at: http://www.gesis.org/eurobarometer-data-service/survey-series/standard-special-eb/sampling-and-fieldwork/

Table 3: Variables and measures in EB 71.3

Identification with Europe Thinking about this, to what extent do you personally feel you are European/ [nationality]? [1 'not at all' – 2 'somewhat' – 3 'not really' – 4 'to a great extent']
Meaning of Europe What does the European Union mean to you personally? Peace; Economic prosperity; Democracy; Social protection; Freedom to travel, study and work anywhere in the EU; Cultural diversity; Stronger say in the world; Euro; Unemployment; Bureaucracy; Waste of money; Loss of our cultural identity; More crime; Not enough control at external borders [0 'no', 1 'yes']
Attitudes towards foreigners Additive index, agreement with positive items/ disagreement with negative items coded 1, agreement to negative items/ disagreement with positive items coded −1, 'it depends' and 'don't know' coded 0. The sum score was divided by the number of valid responses. This results in scaling attitudes from −1 'completely negative' to 1 'completely positive.' QH1_1 People from other ethnic groups enrich the cultural life of >OUR COUNTRY<. + QH1_3 The presence of people from other ethnic groups is a cause of insecurity. – QH1_4 The presence of people from other ethnic groups increases unemployment in >OUR COUNTRY<. – QH1_6 We need immigrants to work in certain sectors of our economy. + QH1_7 The arrival of immigrants in Europe can be effective in solving the problem of Europe's ageing population. + QH1_8 Immigrants can play an important role in developing greater understanding and tolerance with the rest of the world. + QH1_9 Immigrants/legal Immigrants contribute more in taxes than they benefit from health and welfare services. + ['Tend to agree' – 'tend to disagree' – 'it depends' (SPONTANEOUS) – 'don't know']

I will address the content of European identity by studying what the *European Union* means to people. Surveys show that the EU is important for the conception of European identity. In EB 57.2 (European Commission 2012), fielded 2002, the most important reasons for feeling European were 'free movement' (80%) and 'common currency' (76%). Non-EU aspects received considerable but less support ('common history' 54%, 'common ancestry' 52%). Another survey asked young adults across Europe what Europe means to them (Jamieson et al., 2005). 'The EU' (53%) was chosen more often than 'geographical location' (46%) and 'certain values and traditions' (47%). Concentrating on the meaning of the EU focuses on the political aspect of European identity and neglects other meanings. However, the focus is on an aspect people find very relevant.

Method

A simple plot of the dispersion of 14 items on the meaning of the EU will already give a picture of the degree of contestation, i.e., to what extent EU citizens agree or disagree. An exploratory factor analysis (EFA) was the first step to grasp the patterns of the meanings associated with the EU, which is to exploit the fact that some answers go together more often than others. The initial EFA results were used to test a much stricter model that allows distinguishing distinct groups: the latent class factor model (Magidson and Vermunt, 2001). As the term suggests, the latent class factor model bears resemblance to both factor analysis and latent class analysis. Latent factors explain the observed answer patterns; but unlike in a traditional factor analysis, the factors are ordered categorically. Therefore, the latent structure

can be used to cluster observations into segments as in standard latent class analysis. The conversion of categorical factors to clusters is straightforward. For example, consider a model with two dichotomous factors (F1, F2). In this case, the latent class factor model provides probability estimates for the membership of cases in four clusters: cluster 1= (low on F1, low on F2); cluster 2 = (low on F1, high on F2); cluster 3 = (high on F1, low on F2) and cluster 4 = (high on F1, high on F2). Conceptually, this classification is a simple crosstab (though the assignment is probabilistic and not deterministic). Crossing three latent variables with two categories each results in eight clusters. The clustering was used to condition the correlation of identification with Europe and attitudes with foreigners on the meaning of the EU, i.e., to compare the relationships between identification with Europe and attitudes towards foreigners across different types of identity content.

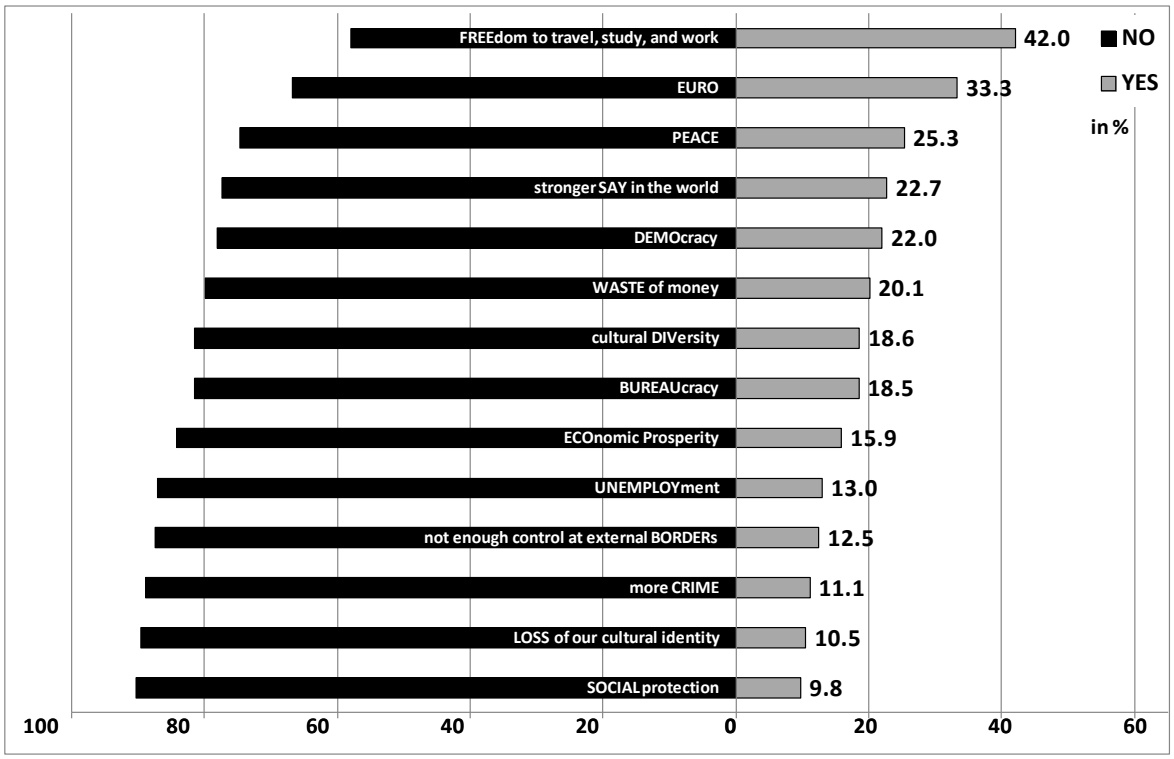
Descriptive results

In the overall sample, the correlation of identification with Europe and attitudes towards foreigners is *r*=.19, which is in line with previous research in two respects: the sign of the association and the rather low size. This is the picture we get when we do not account for the content of European identity; it will serve as a benchmark for the interpretation of the results that do consider the meanings that the respondents attach to Europe.[6]

Figure 2 shows agreement and disagreement with 14 content items. Not a single item is shared by a majority. Equally important, not a single item is unequivocally rejected. Each meaning of the EU is at least endorsed by one out of ten European citizens. The most prevalent meanings are ‘freedom to travel, study, and work in any member state’ (42.0%) and the euro currency (33.3%). But again it is worth noting that even for these two items there is a larger group of citizens who do not think this way. As the case stands now, the degree of contestation is high.

[6] One could argue that the correlation is low because attitudes towards foreigners are multidimensional or not invariant across national contexts (e.g., people may have very different groups in mind when they think about ‘foreigners’). The design of this study compares the low overall correlation with the correlations in subgroups. As long as the groups do not differ in the dimensional structure of attitudes towards foreigners (or the misspecification thereof) and as long as the groups are not highly confounded with national contexts, these concerns do not compromise the goal of this study.

Figure 2: Content and contestation of European identity: The meaning of the EU



Source: EB 71.3, EU27; the short labels used in the following are indicated by capital letters.

Multivariate results

Patterns of meaning and clusters of people

An exploratory factor analysis reveals that three modestly correlated dimensions sufficiently fit the data (see table A1 in the appendix). The first dimension groups the items PEACE, ECO, DEMO, and SOCIAL (see figure 2 for item wording). As the items refer to the democratic state and social welfare, the factor is labeled *social democracy*. The second dimension groups the items FREE, DIV, and EURO. While the commonality of FREE and DIV is straightforward, the alignment of the common currency in this dimension might be puzzling. Apparently, the euro is perceived more in terms of its symbolic quality as a sign of openness or its practical usefulness as a single currency than in regard to its association with economy. The second dimension is labeled *openness & freedom*. Whereas *social democracy* 'looks inward' and stresses collectively shared aspects of the EU as a democratic community, *openness & freedom* 'looks outward' and is about individually enjoyed liberties and the freedom to transgress national boundaries. The third dimension, labeled *the negative*, groups all items that express criticism or discontent, irrespective of their differing content. Models with two negative dimensions, e.g., those separating BUREAU and WASTE from the items that could be seen as representing 'threat' (LOSS, CRIME, BORDER), did not find empirical

support. The item SAY could not be assigned to any dimension and was excluded from further analysis.

The latent class factor model builds on the EFA structure, but uses dichotomous factors. The essential model parameters are the response probabilities for the items that are allowed to differ between the two latent classes of each latent variable. A combination of the positions on the three latent variables should reproduce the response patterns of the individuals on the 13 observed items. To illustrate what a tremendous task the model is supposed to accomplish, note that 13 dichotomous variables amount to 2^{13}=8,192 possible response patterns, 3,153 of which were actually used by respondents, whereas the latent variable pattern of the model has only 2x2x2=8 cells. This sparse model fits the data reasonably well, a likelihood ratio chi^2-test cannot reject it (chi^2=8334, *df*=8147, *p*=.07). The classification quality—i.e., how well the assignment to latent classes can reproduce the observed response patterns—is far from perfect, yet sufficient (relative entropy=.69, Clark, 2010; Ramaswamy et al., 1993).

Table 4: Results of the latent class factor model, estimated probabilities

Latent dimension	**A priori prob.**	**Item**	**P1 (1-1-1) low salience: *empty***	**P2 (1-1-2) medium salience: *the negative***	**P3 (2-1-1) medium salience: *social democracy***	**P4 (2-2-2) high salience: *all-embracing***	**P5 (2-2-1) high salience: *social democracy + openness & freedom***	**P6 (2-1-2)**
1 *Social democracy*	0.26	PEACE	0.12	0.12	0.43	0.43	0.43	0.43
	0.18	ECOPROS	0.06	0.06	0.34	0.34	0.34	0.34
	0.22	DEMO	0.06	0.06	0.44	0.44	0.44	0.44
	0.12	SOCIAL	0.04	0.04	0.23	0.23	0.23	0.23
2 Openness & freedom	0.48	FREE	0.40	0.40	0.40	0.92	0.92	0.40
	0.18	DIV	0.10	0.10	0.10	0.65	0.65	0.10
	0.33	EURO	0.28	0.28	0.28	0.64	0.64	0.28
3 *The negative*	0.14	UNEMPLOY	0.05	0.32	0.05	0.32	0.05	0.32
	0.19	BUREAU	0.09	0.39	0.09	0.39	0.09	0.39
	0.19	WASTE	0.05	0.47	0.05	0.47	0.05	0.47
	0.10	LOSSID	0.03	0.24	0.03	0.24	0.03	0.24
	0.15	CRIME	0.03	0.38	0.03	0.38	0.03	0.38
	0.14	BORDER	0.05	0.30	0.05	0.30	0.05	0.30
Cluster size			29.7%	25.2%	34.1%	3.9%	6.8%	0.2%

Source: EB 71.3, EU27; software: Mplus 6.11, estimator: robust ML;
grey cells: probability >.25 or probability >1.50 times the a priori probability; dark grey cells: probabilities >.50;
P2 (1-1-2) denotes response pattern 2: 1 = in the first latent class on dimension 1 (*social democracy*), 1 = in the first latent class on dimension 2 (*openness & freedom*), 2 = in the second latent class on dimension 3 (*the negative*).
Latent class pattern 1-2-1 and 1-2-2 do not occur;
cluster size: weighted by population size, assignment to classes based on most likely latent class pattern.
For full item labels, see figure 2.

The numbers in table 4 are probabilities—the columns to the left of the items display the probabilities of agreement a priori, which are the observed frequencies of the items[7]; the columns to the right of the items show estimated response probabilities within a latent class pattern. Each combination of latent classes corresponds to a cluster of cases, i.e., to a group of people with a distinct response pattern to the meaning of the EU items. The term 'pattern' stresses latent (class factor) variables and measurement; the terms 'cluster' or 'cross classification' stress types and groups. These are different perspectives on the same model: a certain pattern of latent class variables represents a cluster of cases.

One cluster (P6) is so sparsely populated that statistical uncertainty forbids interpretation and two other cells in the latent variable cross classification are empty, i.e., no cases are assigned to them. For this reason, there remain five substantive types of meaning of the EU. To interpret what is characteristic of a latent class pattern, both the absolute size of the estimated response probability and the ratio of the a priori and the estimated response probability are important. For example, the entry of .92 for the item FREE in the response pattern P4 means that it is almost certain and roughly twice as likely as in the overall sample that someone in cluster 4 will agree that the EU means freedom to travel, study, and work. Therefore, the labeling of the clusters reflects two criteria. The first one is *salience*: To which extent do the respondents assign meaning to the EU? (That is, are there items with a fair chance of affirmative answers?) The second one is *content*: Which of the three content dimensions do they endorse?

P1 groups those 29.7% whose response probabilities are below average on all dimensions. It is very likely that they answered 'no' to every item, except for some realistic chance of agreeing with FREE (0.40). To this group, the EU has no other meaning; its salience is low, the image almost *empty*. For that quarter of respondents in P2 (25.2%), the EU has a predominantly *negative* image. Apart from considerable (but below average) support for FREE and EURO, they blame and criticize the EU. P3, the largest group (34.1%), frames the EU positively as *social democracy* but does not see the EU as particularly standing for *openness & freedom*. In P2 and P3 the salience of the EU is higher than in P1 but lower than in the two remaining clusters. P4 is a minority (3.9%) that attributes meaning to the EU in all three dimensions (*all-embracing*); the EU is highly salient to them, and the picture is mixed.

[7] Differences to the percentages in figure 1 are due to weighting.

P5 (6.7%) comprises people with a salient, unambiguously positive view: the EU stands for *social democracy* and especially for *openness & freedom*.

Who are they? Manifesting the latent classes

Table 5: Characteristics of the latent classes

Cluster	Age: mean (sd)	Education: % higher	Political orientation: mean (sd)	Identification with Europe: mean (sd)	Attitudes towards foreigners
P1*low salience: empty*	44.5 (19.1)	25.4	5.28 (2.06)	1.89 (.93)	–.09 (.57)
P2 *medium salience: the negative*	50.0 (17.5)	21.7	5.36 (2.12)	1.64 (.99)	–.24 (.55)
P3 *medium salience: social democracy*	45.4 (18.8)	29.1	5.29 (2.14)	2.21 (.77)	.02 (.56)
P4 *high salience: all-embracing*	48.0 (17.1)	39.0	5.05 (1.95	2.24 (.77)	.03 (.54)
P5 *high salience: social democracy + openness & freedom*	42.7 (19.6)	45.2	4.90 (1.97)	2.46 (.64)	.26 (.54)
Total = EU27 citizens	46.2 (18.6)	27.5	5.26 (2.10)	2.00 (.91)	–.06 (.57)

Source: EB 71.3, EU27; higher education = left education at the age of 20+; political orientation: left=1, right=10, self-placement; identification with Europe: 1=low, 4 =high; attitudes towards foreigners: –1 totally negative, 1= totally positive; see table 3 for full question wording and index construction.

Table 5 compares the five clusters. The age structure is not very different between the clusters, though in P2, with a negative image of the EU, older people are slightly overrepresented, while they are underrepresented in P5 with a highly salient picture of the EU as *social democracy* and *openness & freedom*. The educational differences are greater: in P4 and P5, the high salience groups, the higher-educated are overrepresented. The share of higher-educated is lowest in the negative cluster, P2, even though it is still one in five who continued education at the age of 20 in that cluster. Politically, all the clusters resemble the general population quite well: on average they are centrists, slightly to the left of the midpoint of the scale at 5.5. P5 is the most leftist cluster. The differences in identification with Europe are more pronounced: P2, the negative image cluster, is almost half a scale point below average; P4, the positive image cluster, is nearly half a scale point above average. This is quite significant on a 4-point scale. Similarly, attitudes towards foreigners are most negative in P2 (–.24) and most positive in P4 (.26). Again, the size of the differences is substantial, given that their scale is –1 to 1. The standard deviations, however, signal considerable variation in attitudes towards foreigners and in identification with Europe in all clusters. In each cluster, there are people with pro- and anti-foreigner sentiments, just as there are people who weakly and who strongly identify with Europe. Hence, there is enough variation within the clusters to merit investigating the relation between identification with Europe and attitudes towards foreigners for each cluster—for if the groups were heavily confounded with these two variables it would not make sense to look at their relation within groups.

Another aspect is the distribution of the clusters within countries. This aspect is interesting in itself, but could in extremis also jeopardize further analyses.

Table 6: Proportions of clusters within countries

Country	**P1** *low salience: empty*	**P2** *medium salience: the negative*	**P3** *medium salience: social democracy*	**P4** *high salience: all-embracing*	**P5** *high salience: social democracy + openness & freedom*	**P6**
FRANCE	25.1	31.4	30.6	4.0	8.4	0.6
GERMANY E	18.0	42.9	22.3	8.6	7.0	1.2
GERMANY W	16.7	31.1	30.6	9.1	12.2	0.2
GREAT BRITAIN	40.6	38.1	17.1	1.6	2.5	0.1

Source: EB 71.3, EU27.

Table 6 shows the proportions of the clusters in France, Germany, and Great Britain—this corresponds to the 'national frames' that Diez Medrano (2003) investigated.[8] In Great Britain, the indifferent cluster, P1, is much larger than in France and Germany. In Eastern Germany, criticism (P2) is much more common than in Western Germany or France. Highly salient images of *social democracy* or *openness & freedom* (P5) are five times more frequent in Western Germany than in Great Britain. We see that our classification also uncovers national differences. However, if we shift our focus away from dissimilarity, we see that the national borders are not identical with the boundaries of the clusters: the different types of meaning associated with the EU are present within each country.

The relation of European identity and xenophobia: a closer look

Finally, we are in a confident position to address the research question. We have found five types of meaning of the EU, i.e., five ways people relate to one of the main objects of European identity, and hence five different types of identity content. Recall that the hypothesis is that the relation of European identity and attitudes towards xenophobia depends on the content of European identity. Technically, this is an interaction model. We expect that the content of European identity will moderate the strengths of the association between identification with Europe and attitudes towards foreigners.

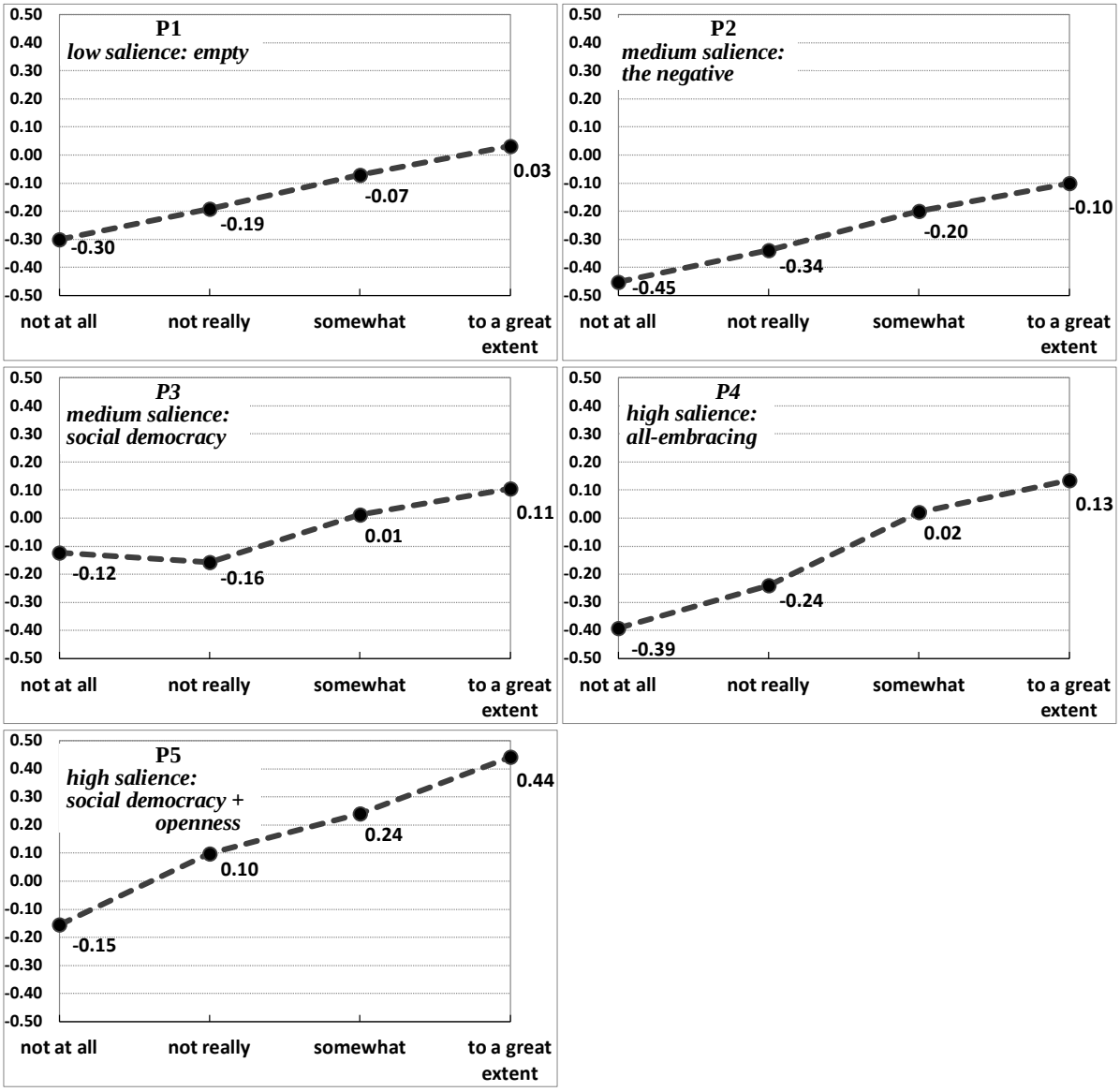
To frame the EU as *openness & freedom* implies to endorse diversity and a positive stance on transgressing national borders and hence should increase tolerance. The effect of *social democracy* is less clear. On the one hand, democratic values can be framed as universal (and thus open to foreigners). On the other hand, social welfare can be seen as a good tied to

[8] A table of the cluster distribution in all EU27 countries can be found in the appendix, table A2.

membership in the nation state (and thus as exclusive). Given this ambiguity, we can, at best, formulate a tentative assumption: the association of identification with Europe and attitudes towards foreigners ought to be lower in P3, the cluster characterized by a modest approval of *social democracy* but not of *openness & freedom*, as compared to P4 and P5. At first sight, hypothesizing on the effect of the *negative* and the *empty* view poses a puzzle. The reasoning on an interaction effect of identification and identity content rests on the assumption that different people identify with Europe to a different degree within each type of identity content. Empirically, we do indeed observe this variation. But why should someone identify with Europe if she thinks that the EU is a bad thing or if he thinks that the EU is meaningless? Two explanations are plausible. One is criticism: stressing the negative aspects could express discontent with the EU as is and sympathy for change, but not the negation of European integration in general. The second is undercoverage of content: respondents can either have affirmative views of the EU that are not reflected in the items or envisage objects other than the EU for their European identity. While these arguments explain how someone in the *empty* or the *negative* cluster can still rank high on identification with Europe, they also reveal that the data remain silent about their reasons for identifying with Europe. Therefore, we have no good reason to speculate on the size of the interaction effect. This argument also suggests that both the *empty* and the *negative* cluster capture the 'true' meaning of the EU or the proper object of European identity only for those who do not identify with Europe ('The EU is negative/meaningless, so I don't identify with Europe'). As the negative items capture a threat to cultural identity and security, we can tentatively assume that those who have negative views of the EU and do not identify with Europe will have the most negative attitudes towards foreigners. This is a weak hypothesis in regard to xenophobia, not an interaction hypothesis.

In sum, the expectations are these: (1) The association of identification with Europe and attitudes towards foreigners is stronger in the clusters P5 and P4 than in the clusters P1, P2, and P3. (2) Attitudes towards foreigners are most negative for the 'non-identifiers' with a negative view of the EU in P2 and P4. As the types of identity content are developed exploratively in this study, I could not a priori state hypotheses on their specific influence on the relationship of identification with Europe and attitudes towards foreigners—the expected ranking of the clusters capitalize on knowing the data. The main hypothesis, however, is theoretical: there are differences in these associations between clusters, and they are substantial.

Figure 3: The relationship of identification with Europe and attitudes towards foreigners within the clusters



Source: EB 71.3, EU27;
x-axis: identification with Europe; y-axis: attitudes towards immigrants, mean values, scale –1 to 1.

There are several ways to describe what figure 3 displays. The description that comes closest to the data is from the perspective of conditional means. The dots in the graphs are the means of attitudes towards foreigners conditioned on identification with Europe and for each cluster. In other words, when moving from the lower left to the upper right of every graph, we can see how the mean values of attitudes towards foreigners change in each cluster in accordance with the respondent's degree of identification with Europe. Another, equally appropriate way of looking at it is to speak of a regression of attitudes on identification, moderated by the identity content. Once we draw a line to connect the mean values, we are in a regression framework. Although the lines in figure 3 are not perfectly linear, they are reasonably close to compare the linear association, i.e., Pearson's *r*, across clusters. In a regression framework the

differences between clusters correspond to an interaction term—the regression coefficient of identification with Europe is allowed to differ between clusters.[9]

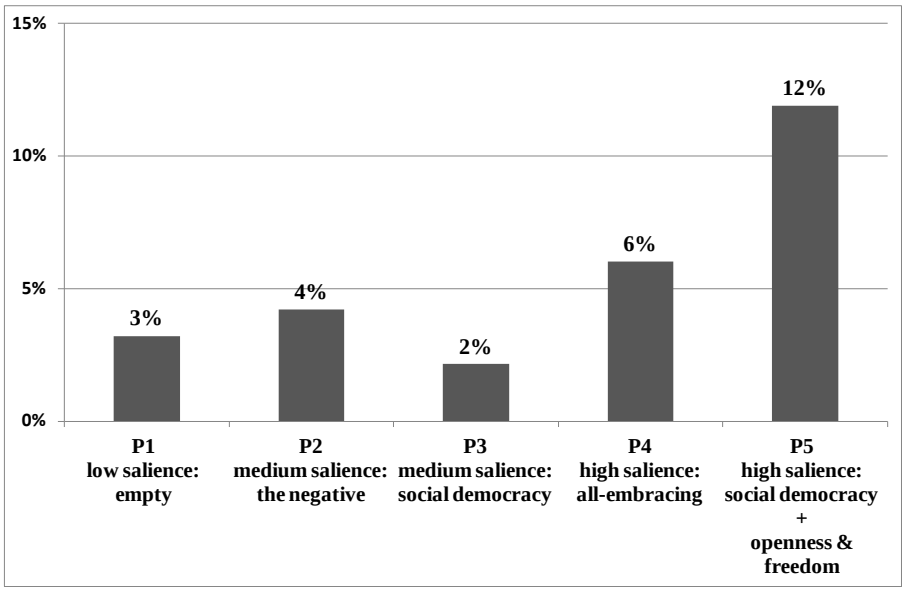
The lower left values in the charts are the levels of xenophobia for those who do not identify with Europe. In line with expectations, the most negative attitudes of non-identifiers occur in P2, the *negative* cluster, and in P4, the *all-embracing* cluster of people who combine *social democracy*, *openness & freedom*, and *negative* views. Mean values of –.45 and –0.39 on a scale from –1 to 1 indicate a strongly negative attitude towards foreigners. Whereas negative attitudes are similarly pronounced for non-identifiers in P1 (–.30), the *empty* cluster, they are less severe in the *social democracy* and/or *openness & freedom* clusters P3 and P5 (–.12 and –.15 respectively).

All regression lines go from the bottom left to the upper right. The association of attitudes towards foreigners and identification is positive in all clusters. What differs is the strength of the association. This pattern of differences supports our assumptions. The correlation is highest in P5 (.34) and second highest in P4 (.24), i.e., among the respondents who emphasize that the EU stands for *openness & freedom*. In the *social democracy* cluster, P3, the correlation is lower (.15). Following the argument that we do not know why people in the empty and in the negative cluster identify with Europe, the positive associations in P1 (.18) and P2 (.20) cannot be interpreted in terms of content—they are yet another hint at undercoverage of content.

Ultimately, the correlations are positive across the board, so why make such a big fuss of it? The correlations range from 0.15 to 0.34—the highest number being roughly two times larger than the smallest number. Correlation coefficients are misleading if we look at them this way. A true-to-scale picture of association is the amount of shared variance, which is not r but r^2. This changes the picture.

[9] I abstain from reporting and relying on statistical significance for the interaction effect and any other differences between clusters. High sample size leads to high power, i.e., test statistics detect very small deviations from traditional null hypotheses (for a thorough discussion of the assumptions and pitfalls of null hypothesis significance testing, see, e.g., Cohen, 1994). Were we to rely on *p*-values alone, we would find the difference between r=.18 in P1 and r=.15 P3 to be highly significant. All differences turn out statistically significant if the sample is large enough. From a substantive perspective, however, only some of them are interesting. If average standardized effects in previous research are around 0.10, I suggest that differences greater than .05 (i.e., half the size of the average effect) are noteworthy.

Figure 4: Variance overlap of identification with Europe and attitudes towards foreigners in the five clusters



Source: EB 71.3, EU27.

Figure 4 compares across clusters how much variance attitudes towards foreigners share with identification with Europe. In cluster P5, identification with Europe and attitudes towards foreigners share 12% of their variance, while in P3, it is only 2%. In terms of shared variance the ratio is not 1:2 but rather 1:6. This makes a difference.

Limitations

Though some limitations are specific to this study, the underlying issues—the selection of the object of identification and the selection of the items to tap its meaning—are generic to any quantitative study of identities. This study is restricted to the EU as the object of European identity—for good reasons: the EU is relevant to many people and it is clear what the EU is. Using the vague term 'Europe' instead could even be an asset if, and only if, a study could work with a larger set of items to explore the facets of its meaning. Other potential objects such as '(European) culture' or '(European) values' are less specific than they might seem, i.e., they demand larger sets of items as well.

Even for the EU, the listing of meanings in the EB lacks some of the 'usual suspects.' There are no items on 'social Europe' in the sense of social justice and on social or ecological sustainability. A hint at omitted meanings are the people in the *empty* cluster who still identify with Europe. The most plausible interpretation here is that Europe has a meaning to them other then the meanings prompted. The better the range of potential meanings are grasped, the better interpretable and the more pronounced the differences in the correlations should be.

Had there been positively worded items on ‘euro-chauvinism’ or ‘a Europe of fatherlands’, we can assume that even a negative association of identification with Europe and attitudes towards foreigners could have been discovered for a subgroup. Granted, this is mere speculation.

Apart from the specific issues in this study, any list of meanings will always be incomplete to someone. Incompleteness is a problem generic to surveys on meanings. Other data sources with other measures would more likely than not lead to somewhat different configurations of meanings—especially for very contested identities.

Conclusions

In a nutshell, the point is that content counts! Content counts for the study of European identity and for the study of any identity.

European identity is associated with hopes—not the least of which is that European identity would do without the xenophobic drawbacks of national identity. So far, evidence whether or not European identity lives up to this ideal is inconclusive. I have argued that this is not surprising given how we try to measure European identity in surveys. Social theory of collective identities stresses content and contestation. But empirical social science usually treats European identity as a one-dimensional variable and uses measures of identification with Europe only. These measures include neither content nor contestation. Social scientists are not blind to this gap; they feel uneasy but short of any better alternative. Reflecting on the Eurobarometer, Díez Medrano and Gutiérrez (2001, 757) called researchers ‘prisoners of the limitations of some of the indicators used in their source of data of choice.’

This dilemma cannot be resolved, but it can be addressed. The endeavor is to bring content in. While some researchers simply echo the problematic dichotomy of ‘good’ patriotism and ‘bad’ nationalism, I have proposed an exploratory approach to tap an important aspect of the content of European identity: the meaning of the EU. A latent class factor model summarizes 13 items into three dimensions of meaning that cluster the EU citizens into five groups. Although positive in all clusters, the strength of the relation of identification with Europe and attitudes towards foreigners differs substantially between the clusters. Therefore, the answer to the question whether or not European identity is a safeguard against xenophobia is: ‘It depends!’ —it depends on the content of European identity. If the image of the EU is *openness & freedom*, identification with Europe indeed comes with more tolerant attitudes. If the EU is mainly seen as *social democracy* or has a *negative* image, this relation is much weaker.

To sharpen the argument, if the content of European identity is tolerance, then the people who identify with Europe more strongly are more tolerant. This is close to tautological—and disappointing, but only if we think of it in causal terms. European identity does not lead to or cause anything automatically, rather its content can be affine to something or not. European identity is no fast lane to tolerance. Research should not treat European identity as a (single) variable, but shift the question from 'Is there a European identity?' to 'Which European identities are there?' For policy, the task is not to foster just any European identity but a specific one. European identity is no cure in itself, but a European identity that means openness and tolerance may act as remedy to xenophobia. European identity is a good example to demonstrate the importance of identity content because contestation is high, as there are very different opinions what Europe ought to be like. But if we buy into the argument that—in principle—all collective identities are about contested content, we cannot but agree that identities are socially constructed. Social constructivism is a red rag to many realist, empirical social scientists. For some variants of constructivism and some variants of realism, the gap may indeed be real and irreconcilable. This study wants to make a case that basic constructivist thinking and quantitative analyses can go together. This is not an easy task, but in my view a necessary one and well worth taking up the challenge.

For the study of identities, a basic social constructivist account—by which I mean a focus on meaning, construction, and contestation—seems more appropriate than any other, potentially essentialist theoretical orientation. Identities are about shared stories, and research cannot get at identities without reconstructing the different meanings people associate with these identities. In this endeavor, social constructivism prevents us from searching for the 'true' core of any identity and rather directs us towards alternative constructions.

That said, should I now be tempted to stop quantitative work on identities and defect to the qualitative camp with banners flying? While fieldwork and qualitative analyses are the tools of choice to reconstruct identity content and contestation, they are limited in scope and problematic in terms of selection. Quantitative science can summarize more information and weigh the information equally. Hence, there is no need to abandon one for the other but to (1) acknowledge the epistemological strengths and limits of either approach and (2) push hard to incorporate appropriate theory in our investigations.

I, for my part, will stick to the camp that I know best: quantitative social research. Employing numbers, however, comes at a price. It is pointless to accuse survey analyses of loosing individual perspectives. We are not interested in rare cases. In the same vein, it is unfruitful to

argue that surveys fail to capture each and every meaning associated with the EU or Europe. They do not have to. Surveys provide an average picture. But a built-in neglect for individuality is no carte blanche that justifies getting the average picture completely wrong. Survey research does not have to grasp all meanings of an identity, but we should grasp the most important ones. This task is still tedious enough. We need to decide which meanings to survey, and it is obvious that fieldwork and pretesting are beneficial to make these decisions. We also need to allow for change of meaning, which is to say that we need to allow for change of measures as well. This is cumbersome because we may lose some of the neat characteristics of variables that we like in our models. We can neither theoretically assume nor should we empirically strive for constructs that remain the same across time and social contexts. The goal is to tap the meanings that are out there now (and not only the meanings we asked about last time). This does not render comparison impossible. Some identity content may be longstanding and wide-ranging and thus comparable, whereas other identity content may be cursory. This is exactly the empirical question of interest. Only if we can incorporate the meanings that constitute an identity into our models, can we meaningfully address the question of its association with other social phenomena.

References

Abdelal, R., Herrera, Y. M., Johnston, A. I. & McDermott, R. (2006) Identity as a Variable. *Perspectives on Politics* 4: 695-711.

Bauman Z (2000) *Liquid Modernity.* Cambridge: Polity Press.

Beck, U. & Grande, E. (2004) *Das kosmopolitische Europa. Gesellschaft und Politik in der zweiten Moderne.* Frankfurt am Main: Suhrkamp.

Billig, M. (1995) Rhetorical psychology, ideological thinking, and imagining nationhood. In Johnston H and Klandermans B (eds.) *Social Movements and Culture*. Minneapolis: University of Minnesota Press.

Blank, T. & Schmidt, P.(2003) National Identity in a United Germany: Nationalism or Patriotism? An Empirical Test With Representative Data. *Political Psychology* 24:289-312.

Brubaker, R. (1994) *Citizenship and Nationhood in France and Germany.* Cambridge: Harvard University Press.

Brubaker, R. (1999) The Manichean myth: Rethinking the distinction between 'civic' and 'ethnic' nationalism. In: Kriesi H (ed.) *Nation and National Identity. The European Experience in Perspective.* Chur: Rüegger, 272–305.

Brubaker, R. & Cooper, F. (2000) Beyond 'identity'. *Theory and Society* 29: 1–47.

Bruter, M. (2004) Civic and cultural components of a European identity: A pilot model of measurement of citizens' level of European Identity. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield 186–213.

Bruter, M. (2005) *Citizens of Europe? The Emergence of a Mass European Identity.* Basingstoke: Palgrave Macmillan.

Cerutti, F. & Lucarelli, S. (eds.) (2008) *The Search for a European Identity. Values, Policies and Legitimacy of the European Union.* London: Routledge.

Checkel, J. T. & Katzenstein, P. J. (2009) *European Identity.* Cambridge: Cambridge University Press.

Citrin, J. & Sides, J. (2004) More than nationals: How identity choice matters in the New Europe. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield 161–185.

Clark, S. L. (2010) Mixture Modeling with Behavioral Data. Doctoral dissertation, University of California, Los Angeles.

Dejaeghere, Y. & Quintelier, E. (2008) Does European Citizenship Increase Tolerance in Young People? *European Union Politics* 9:339–362.

Delanty, G. & Rumford. C. (2005) *Rethinking Europe. Social theory and the implications of Europeanization.* London: Routledge.

Díez Medrano, J. (2003) *Framing Europe. Attitudes to European Integration in Germany, Spain, and the United Kingdom.* Princeton: Princeton University Press.

Díez Medrano, J & Gutiérrez, P. (2001) Nested identities: National and European identity in Spain. *Ethnic and Racial Studies* 24:753–778.

European Commission (2012) Eurobarometer 71.3 (Jun-Jul 2009). TNS OPINION & SOCIAL, Brussels. GESIS Data Archive, Cologne. ZA4973 datafile version 3.0.0, doi:10.4232/1.11135.

European Commission (2012) Eurobarometer 57.2 (Apr-Jun 2002). European Opinion Research Group, Brussels. GESIS Data Archive, Cologne. ZA3641 data file version 1.0.1, doi:10.4232/1.10951.

Fearon, J. D. (1999) *What Is Identity (As We Now Use the Word)?* Stanford: Stanford University.

Fligstein, N. (2008) *Euroclash. The EU, European Identity, and the Future of Europe.* Oxford: Oxford University Press.

Fuß, D. (2006) *Europa als Quelle sozialer Identität*. Dissertation, University of Gießen, Germany.

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A. & Rust M. C. (1993) The Common Ingroup Identity Model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology* 4: 1–26.

Gerhards, J. (2003) Identifikation mit Europa. Einige begriffliche Vorklärungen. In: Allmendinger, J. (ed.) *Entstaatlichung und soziale Sicherheit.* Opladen: Leske + Budrich, 467–474.

Graham, J. W. (2009) Missing data analysis: Making it work in the real world. *Annual Review of Psychology* 60:549–576.

Habermas, J. (1998) *Die postnationale Konstellation. Politische Essays*. Frankfurt am Main: Suhrkamp.

Habermas, J. & Derrida, J. (2003) Unsere Erneuerung. Nach dem Krieg: Die Wiedergeburt Europas. *FAZ*, 31 May, 33–34.

Heyder, A. & Schmidt, P. (2002) Deutscher Stolz. Patriotismus wäre besser. In: Heitmeyer, W. (ed.): *Deutsche Zustände. Folge I*. Frankfurt am Main: Suhrkamp, 71–82.

Hobsbawm, E. J. (1992) *Nations and Nationalism since 1780. Programme, Myth, Reality.* Cambridge: Cambridge University Press.

Huddy, L. (2001) From social to political identity: A critical examination of Social Identity Theory. *Political Psychology* 22:127–156.

Jenkins, R. (1996) *Social Identity*. London: Routledge.

Kosterman, R. & Feshbach, S. (1989) Toward a measure of patriotic and nationalistic attitudes. *Political Psychology* 10:257–274.

Licata, L., & Klein, O. (2002) Does European citizenship breed xenophobia? European identification as a predictor of intolerance towards immigrants. *Journal of Community and Applied Social Psychology* 12:1-15.

Magidson, J. & Vermunt, J. K. (2001) Latent class factor and cluster models, bi-Plots, and related graphical displays. *Sociological Methodology*. 31:223–264.

McGarty, C. (2001) Social Identity Theory does not maintain that identification produces bias, and self-categorization Theory does not maintain that salience is identification: Two comments on Mummendey, Klink and Brown. *British Journal of Social Psychology*, 40:173-176.

Mühler, K. & Opp, K. (2004) *Region und Nation. Zu den Ursachen und Wirkungen regionaler und überregionaler Identifikation.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Mummendey, A. & Waldzus, S. (2004) National Difference and European Plurality: Discrimination or Tolerance between European Countries. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 59-72.

Opp, K. (2003) Die Identifikation der Bürger mit Europa. Einige Hypothesen und ihre empirische Überprüfung. In: Allmendinger, J. (ed.) *Entstaatlichung und soziale Sicherheit.* Opladen: Leske + Budrich, 475–497.

Pichler, F. (2008) How Real is cosmopolitanism in Europe? *Sociology* 44:1107–1126.

Ramaswamy, V., DeSarbo, W., Reibstein, D. & Robinson, W. (1993) An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science* 12:103–124.

Reicher, S. (2004) The Context of Social Identity: Domination, Resistance, and Change. *Political Psychology*, 25:921–945.

Reicher, S. & Hopkins, N. (2001) *Self and Nation.* London: Sage

Schatz, R. T., Staub, E. & Lavine, H. (1999) On the varieties of national attachment: Blind versus constructive patriotism. *Political Psychology* 20:151–174.

Skrobanek, J. (2004) Soziale Identität und Ausländerfeindlichkeit. *Berliner Journal für Soziologie* 14: 357–377.

Tajfel, H. & Turner, J. C. (1986) The social identity theory of intergroup behavior. In: Austin, W. G. & Worchel, S. (eds.) *Psychology of Intergroup Relations.* Chicago: Nelson-Hall Publishers, 7–24.

Tilly, C (2003) Political identities in changing polities. *Social Research* 70:605–620.

Waldzus, S., Mummendey, A., Wenzel, M. & Weber, U. (2003) Towards tolerance: Representations of superordinate categories and perceived ingroup prototypicality. *Journal of Experimental Social Psychology* 39:31–47.

Weiler, J. H. H. (1999) *The Constitution of Europe. Essays on the Ends and Means of European Integration.* Cambridge: Cambridge University Press.

Wimmer, A. (2002) *Nationalist Exclusion and Ethnic Conflict. Shadows of Modernity.* Cambridge: Cambridge University Press.

Yack, B (1998) Can patriotism save us from nationalism? Rejoinder to Viroli. *Critical Review* 12:203–206.

Appendix

Table A1: Exploratory factor analysis, standardized factor loadings

	Factor 1 *social democracy*	**Factor 2** *openness & freedom*	**Factor 3** *the negative*
PEACE	**0.37**	0.28	0.07
ECOPROS	**0.53**	0.04	–0.07
DEMO	**0.47**	0.25	0.00
SOCIAL	**0.67**	–0.01	0.15
FREE	0.00	**0.57**	–0.17
DIV	0.19	**0.57**	0.01
SAY	0.26	0.27	–0.07
EURO	0.03	**0.36**	0.09
UNEMPLOY	0.02	–0.21	**0.64**
BUREAU	–0.30	0.27	**0.39**
WASTE	–0.36	0.00	**0.54**
LOSSID	0.14	–0.07	**0.50**
CRIME	0.02	0.03	**0.77**

Source: EB 71.3, EU27;
Estimator: WLSMV, Geomin rotation (factor correlations: r_{12}=0.23, r_{13}=-0.20, r_{23}=0.12);
Model fit: chi^2=1263.4, *df*=52 , *RMSEA*=0.029, *pclose*=1, *CFI*=0.951.

Table A2: Proportions of clusters within all EU27 Countries

Country	Latent class pattern = cluster					
	P1	P2	P3	P4	P5	P6
AUSTRIA	14.3	43.6	20.0	16.2	4.9	1.1
BELGIUM	23.1	27.7	33.8	6.1	9.2	0.2
BULGARIA	27.9	9.7	52.0	1.1	9.2	0.0
CYPRUS (REP.)	9.5	26.2	38.1	19.0	7.1	0.0
CZECH REPUBLIC	29.6	30.5	31.5	4.5	3.7	0.2
DENMARK	17.2	31.3	29.3	10.4	11.8	0.0
ESTONIA	40.0	15.0	31.7	3.3	10.0	0.0
FINLAND	30.6	30.2	27.1	3.4	8.6	0.0
FRANCE	25.1	31.4	30.6	4.0	8.4	0.6
GERMANY E	18.0	42.9	22.3	8.6	7.0	1.2
GERMANY W	16.7	31.1	30.6	9.1	12.2	0.2
GREAT BRITAIN	40.6	38.1	17.1	1.6	2.5	0.1
GREECE	17.8	31.7	36.7	6.9	6.0	0.9
HUNGARY	34.8	24.4	34.5	1.1	5.0	0.2
IRELAND	31.4	16.8	43.8	1.8	6.2	0.0
ITALY	37.1	18.5	39.0	0.7	4.6	0.0
LATVIA	49.5	21.6	24.7	1.0	3.1	0.0
LITUANIA	33.0	18.3	39.8	1.0	7.9	0.0
LUXEMBOURG	18.5	18.5	37.0	11.1	14.8	0.0
MALTA	30.4	13.0	43.5	4.3	8.7	0.0
NETHERLANDS	27.1	20.7	40.3	4.7	7.1	0.1
NORTHERN IRELAND	47.4	24.2	23.2	0.0	5.3	0.0
POLAND	44.1	12.6	37.6	0.9	4.7	0.0
PORTUGAL	39.2	20.0	32.2	3.7	5.0	0.0
ROMANIA	21.3	11.8	57.8	2.0	7.1	0.0
SLOVAKIA	23.6	21.9	39.9	6.3	8.0	0.3
SLOVENIA	26.1	32.2	28.7	6.1	7.0	0.0
SPAIN	34.2	10.3	46.6	1.2	7.6	0.1
SWEDEN	15.0	34.8	35.0	7.1	8.1	0.0
TOTAL	29.7	25.2	34.1	3.9	6.8	0.2

Source: EB 71.3, EU27.

Bibliography

Abdelal, R., Herrera, Y. M., Johnston, A. I. & McDermott, R. (2006) Identity as a Variable. *Perspectives on Politics* 4:695-711.

Abdelal, R., Herrera, Y. M., Johnston, A. I. & McDermott, R. (eds.) (2009) *Measuring identity. A guide for social scientists.* Cambridge, New York: Cambridge University Press.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999) *Standards for Educational and Psychological Testing. 2nd ed.* Washington DC: American Psychological Association.

Arbuckle, J. L. (2005) *Amos 6.0 User's Guide*. Chicago, IL: SPSS.

Barnea, M. (2003) *Personal values and party orientations in different cultures*. Doctoral dissertation, The Hebrew University of Jerusalem, Israel.

Bartholomew, D. J. (1987) *Latent variable models and factor analysis*. London: Griffin.

Bauman Z (2000) *Liquid Modernity.* Cambridge: Polity Press.

Beck, U. & Grande, E. (2004) *Das kosmopolitische Europa. Gesellschaft und Politik in der zweiten Moderne.* Frankfurt am Main: Suhrkamp.

Becker, H. S. (1967) Whose side are we on? *Social Problems* 14:239–247.

Beckers, T., Siegers, P., & Kuntz, A. (2012) Congruence and performance of value concepts in social research. *Survey Research Methods* 6:13-24.

Beierlein, C., Davidov, E., Schmidt, P., Schwartz, S. H., & Rammstedt, B. (2012) Testing the discriminant validity of Schwartz Portrait Value Questionnaire items- A replication and extension of Knoppen and Saris. *Survey Research Methods* 6:25-36.

Bezonsky, M. D., Ciecuch, J., Durieux B., & Soenen, B. (2011) The how and what of identity formation: Associations between identity styles and value orientations. *Personality and Individual Differences* 50:295-299.

Billiet, J. (2003) Cross-cultural Equivalence with structural equation modeling. In: Harkness, J. A., Van de Vijver,F. & Mohler, P. P. (eds.) *Cross-cultural survey methods.* New York: Wiley, 247-264.

Billig, M. (1995) Rhetorical psychology, ideological thinking, and imagining nationhood. In Johnston H and Klandermans B (eds.) *Social Movements and Culture*. Minneapolis: University of Minnesota Press.

Bilsky, W. & Schwartz, S. H. (1994) Values and personality. European Journal of Personality 8, 163-181.

Blank, T. & Schmidt, P.(2003) National Identity in a United Germany: Nationalism or Patriotism? An Empirical Test With Representative Data. *Political Psychology* 24:289-312.

Block, N. J. (1976) Fictionalism, Functionalism and Factor Analysis. In: Buck, R. C., , Schaffner, K. F. & Suppe, F (eds.) *Synthese library, vol. 101, Proceedings of the biennial meeting of the Philosophy of Science Association*. Dordrecht: Reidel.

Bollen, K. & Lennox, R. (1991) Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin* 110:305–314.

Bollen, K. A. & Bauldry, S. (2011) Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods* 16:265–284.

Bollen, K. A. (1989) *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A. (2002) Latent variables in psychology and the social sciences. *Annual review of psychology* 53:605–634.

Borsboom, D. (2005) *Measuring the mind. Conceptual issues in modern psychometrics.* Cambridge : Cambridge University Press.

Borsboom, D. (2006) The attack of the psychometricians. *Psychometrika* 71:425–440.

Borsboom, D., Mellenbergh, G. J. 6 van Heerden, J. (2003) The theoretical status of latent variables. *Psychological Review* 110:203–219.

Borsboom, D.,Cramer, A. O. J. , Kievit, R. A., Zand Scholten, A., & Franić, S. (2009) The End of Construct Validity., In: , Lissitz, R. W (ed.) *The concept of validity. Revisions, new directions, and applications.* Charlotte, NC: Information Age Publishing, pp. 135–170.

Bridgman, P. W. (1927) *The logic of modern physics.* New York: Macmillan.

Brubaker, R. & Cooper, F. (2000) Beyond 'identity'. *Theory and Society* 29: 1–47.

Brubaker, R. (1994) *Citizenship and Nationhood in France and Germany.* Cambridge: Harvard University Press.

Brubaker, R. (1999) The Manichean myth: Rethinking the distinction between 'civic' and 'ethnic' nationalism. In: Kriesi H (ed.) *Nation and National Identity. The European Experience in Perspective.* Chur: Rüegger, 272–305.

Bruter, M. (2004) Civic and cultural components of a European identity: A pilot model of measurement of citizens' level of European Identity. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield 186–213.

Bruter, M. (2005) *Citizens of Europe? The Emergence of a Mass European Identity.* Basingstoke: Palgrave Macmillan.

Byrne, B. & van de Vijver, F.(2010) Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing* 10:107–132.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989) Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin* 105: 456-466.

Campbell, D. T. & Fiske, D. W. (1959) Convergent and discriminant validation by the multitrait multimethod matrices. *Psychological Bulletin* 56:81-105.

Caprara, G. V., Schwartz, S. H., Capanna, C., Vecchione, M., & Barbaranelli, C. (2006) Personality and politics: Values, traits, and political choice. *Political Psychology*27:1-28.

Carnap, R. (1936) Testability and meaning I. *Philosophy of Science* 3:419–471.

Carnap, R. (1937) Testability and meaning II. *Philosophy of Science* 4:1–40.

Cerutti, F. & Lucarelli, S. (eds.) (2008) *The Search for a European Identity. Values, Policies and Legitimacy of the European Union.* London: Routledge.

Checkel, J. T. & Katzenstein, P. J. (2009) *European Identity.* Cambridge: Cambridge University Press.

Chen, F. F. (2007) Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling* 14:464-504.

Cheung, G. W. & Rensvold, R. B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling* 9: 233-255.

Chou, C.-P., Bentler, P. M., & Satorra, A. (1991) Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study. British Journal of Mathematical and Statistical Psychology, 44, 347–357.

Cieciuch, J., Davidov E., Vecchione, M., and Schwartz, Sh. H. (2014) A Hierarchical Structure of Basic Human Values in a Third-Order Confirmatory Factor Analysis. *Swiss Journal of Psychology* 73:177–182.

Citrin, J. & Sides, J. (2004) More than nationals: How identity choice matters in the New Europe. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield 161–185.

Clark, S. L. (2010) Mixture Modeling with Behavioral Data. Doctoral dissertation, University of California, Los Angeles.

Cliff, N. (1983) Some Cautions Concerning The Application Of Causal Modeling Methods. *Multivariate Behavioral Research* 18:115–126.

Coenders, G. & Saris, W. E. (1995) Categorization and measurement quality. The choice between Pearson and polychoric Correlations. In Saris, W. E. & Münnich, Á. (eds.) *The multitrait-multimethod approach to evaluate measurement instruments.* Budapest: Eötvös University Press, 125-144.

Coenders, G., Satorra, A., & Saris, W. E. (1997) Alternative approaches to structural modeling of ordinal data. A Monte Carlo study. *Structural Equation Modeling* 4:261-282.

Coenders, M., Lubbers, M. & Scheepers, P. (2003) Majority populations' attitudes towards migrants and minorities. Vienna: European Monitoring Centre on Racism and Xenophobia (EUMC), Ref. no. 2003/04/01.

Commission of the European Communities (2006) White Paper on a European Communication Policy. Brussels: COM (2006) 35 final.

Costner H. L. & Leik R.K. (1964) Deductions from axiomatic theory. *American Sociological Review* 24:819-835.

Crano, W. D. & Brewer, M. B (2002) *Principles and methods of social research*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Cronbach, L. J. & Meehl, P. E. (1955) Construct validity in psychological tests. *Psychological Bulletin* 52: 281–302.

Davidov, E. & Meuleman, B. (2012) Explaining attitudes towards immigration policies in European countries: The role of human values. *Journal of Ethnic and Migration Studies* 38:757-775.

Davidov, E. & Schmidt, P. (2007) Are values in the Benelux comparable? In: Loosveldt, G., Swyngedouw. M. & Cambré, B. (eds.) *Measuring meaningful data in social research*. Leuven: Acco, 373-386.

Davidov, E. & Schmidt, P. (2007) Are values in the Benelux comparable? In Loosveldt, G., Swyngedouw, M. & Cambré, B. (eds.) *Measuring meaningful data in social research.* Leuven, Belgium: Acco, 373-386.

Davidov, E. (2008) A cross-country and cross-time comparison of the human values measurements with the second round of the European Social Survey. *Survey Research Methods* 2:33-46.

Davidov, E. (2008) A cross-country and cross-time comparison of the human values measurements with the second round of the European Social Survey. *Survey Research Methods* 2:33-46.

Davidov, E. (2010) Testing for comparability of human values across countries and time with the third round of the European Social Survey. *International Journal of Comparative Sociology* 51:171-191.

Davidov, E., Meuleman, B., Billiet, J., & Schmidt, P. (2008b) Values and support for immigration. A cross-country comparison. *European Sociological Review* 24:583-599.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014) Measurement Equivalence in Cross-National Research. *Annual Review of Sociology* 40:55–75.

Davidov, E., Schmidt, P., & Schwartz, S. H. (2008) Bringing values back in: Testing the adequacy of the European Social Survey to measure values in 20 Countries. *Public Opinion Quarterly* 72, 420-445.

De Beuckelaer, A. (2005) *Measurement invariance issues in international management research.* Doctoral dissertation, Hasselt University, Diepenbeek, Belgium.

De Graaf, N. D., Hagenaars, J., & Luijkx, R. (1989) Intragenerational stability of postmaterialism in Germany, the Netherlands and the United States. *European Sociological Review* 5:183-201.

Dejaeghere, Y. & Quintelier, E. (2008) Does European Citizenship Increase Tolerance in Young People? *European Union Politics* 9:339–362.

Delanty, G. & Rumford. C. (2005) *Rethinking Europe. Social theory and the implications of Europeanization.* London: Routledge.

Delanty, G. & Rumford. C. (2005) *Rethinking Europe. Social theory and the implications of Europeanization.* London: Routledge.

Delanty, G. (2005) The Idea of a Cosmopolitan Europe. *International Review of Sociology* 15: 405-421.

Devitt, M. (1997) *Realism and truth*. Princeton: Princeton University Press.

Diamantopoulos, A., Riefler, P., & Roth, K. (2008) Advancing formative measurement models. *Journal of Business Research* 61:1203–1218.

Díez Medrano, J. & Gutiérrez, P. (2001) Nested identities: national and European identity in Spain. *Ethnic and Racial Studies* 24:753–778.

Díez Medrano, J. (2003) *Framing Europe*. *Attitudes to European Integration in Germany, Spain, and the United Kingdom.* Princeton: Princeton University Press.

Díez Medrano, J. (2010) Unpacking European Identity. *Politique Européenne* 30:45–66.

Duncan, O. D. (1984) *Notes on social measurement. Historical and critical.* New York: Russell Sage Foundation.

Duriez, B., Van Hiel, A., & Kossowska, M (2005) Authoritarianism and Social Dominance in Western and Eastern Europe: The Importance of the Sociopolitical Context and of Political Interest and Involvement. *Political Psychology* 26:299-320.

Eagly, A. H. & Chaiken, S. (1993) *The Psychology of Attitudes*. Wadsworth, Fort Worth.

Eriksen, T. H (2002) *Ethnicity and Nationalism: Anthropological Perspectives*. London: Pluto Press.

European Commission (2012) Eurobarometer 57.2 (Apr-Jun 2002). European Opinion Research Group, Brussels. GESIS Data Archive, Cologne. ZA3641 data file version 1.0.1, doi:10.4232/1.10951.

European Commission (2012) Eurobarometer 71.3 (Jun-Jul 2009). TNS OPINION & SOCIAL, Brussels. GESIS Data Archive, Cologne. ZA4973 datafile version 3.0.0, doi:10.4232/1.11135.

European Social Survey (2004) Round 2 Data, data file edition 3.2., 2011. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

Fearon, J. D. (1999) *What Is Identity (As We Now Use the Word)?* Stanford: Stanford University.

Fishbein, M. & Ajzen, I. (2010) *Predicting and Changing Behavior: The Reasoned Action Approach*. New York: Psychology Press.

Flanagan, S. (1987) Changing values in advanced industrial societies revisited: Towards a resolution of the values debate. *American Political Science Review* 81:1303-1319.

Fligstein, N. (2008) *Euroclash. The EU, European Identity, and the Future of Europe.* Oxford: Oxford University Press.

Flora, D. B. & Curran, P. J. (2004) An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* 9:466-491.

Franco, F. (2010) Speech to the Council of Europe. retrieved from: http://www.coe.int/t/dc/files/pa_session/jan_2010/20100126_disc_frattini_en.asp [29.06.2011]

Fuß, D. (2006) *Europa als Quelle sozialer Identität*. Dissertation, University of Gießen, Germany.

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A. & Rust M. C. (1993) The Common Ingroup Identity Model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology* 4: 1–26.

Garrett, H. E. (1937) *Statistics in psychology and education*. New York: Longmans, Green.

Gerhards, J. (2003) Identifikation mit Europa. Einige begriffliche Vorklärungen. In: Allmendinger, J. (ed.) *Entstaatlichung und soziale Sicherheit.* Opladen: Leske + Budrich, 467–474.

Giddens, A. (1984) *The constitution of society. Outline of the theory of structuration*. Cambridge [Cambridgeshire]: Polity Press.

Glen, N. D. (1974) Aging and conservatism. *Annuals of the American Academy of Political and Social Science* 415:176–186.

Gould, S. J. (1996*) The mismeasure of man*. New York: Norton.

Gouldner, A. W. (1962) Anti-Minotaur: The Myth of a Value-Free Sociology. *Social Problems* 9:199-213.

Graham, J. W. (2009) Missing data analysis: Making it work in the real world. *Annual Review of Psychology* 60:549–576.

Guttman, L. (1971) Measurement as structural theory. *Psychometrika* 36:329–347.

Habermas, J. & Derrida, J. (2003) Unsere Erneuerung. Nach dem Krieg: Die Wiedergeburt Europas. *FAZ*, 31 May, 33–34.

Habermas, J. (1998) *Die postnationale Konstellation. Politische Essays*. Frankfurt am Main: Suhrkamp.

Hacking, I. (2004) *Historical ontol*ogy. Cambridge, Mass.: Harvard University Press.

Haig, B. D. (2013) The Philosophy of Quantitative Methods. In: Little, T. D. (ed.) *The Oxford handbook of quantitative methods*. Oxford [u.a.]: Oxford Univ. Press, 7–31.

Hempel, C. G. (1952) *Fundamentals of concept formation in empirical science.* Chicago: University of Chicago Press.

Hermann, R. K. & Brewer. M. B. (2004) Why Study Identity and Institutions? In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 247-271.

Heyder, A. & Schmidt, P. (2002) Deutscher Stolz. Patriotismus wäre besser. In: Heitmeyer, W. (ed.): *Deutsche Zustände. Folge I*. Frankfurt am Main: Suhrkamp, 71–82.

Hobsbawm, E. J. (1992) *Nations and Nationalism since 1780. Programme, Myth, Reality*. Cambridge: Cambridge University Press.

Hofstede, G. (2001) *Culture's Consequences. Comparing Values, Behaviors, Institutions, and Organizations across Nations*. Thousand Oaks: Sage.

Hood, B.S. (2008) *Latent Variable Realism in Psychometrics.* Doctoral Dissertation, Indiana University, United States of America.

Hu, L. & Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* 6, 1-55.

Huddy, L. (2001) From social to political identity: A critical examination of Social Identity Theory. *Political Psychology* 22:127–156.

Inglehart, R. & Baker, W. E. (2000) Modernization, Cultural Change, and the Persistence of Traditional Values. *American Sociological Review* 65:19-51.

Inglehart, R. & Welzel, C. (2005) *Modernization, Cultural Change, and Democracy. The Human Development Sequence*. Cambridge: Cambridge University Press,.

Inglehart, R. (1977) *The Silent Revolution: Changing Values and Political Styles among Western Publics.* Princeton: Princeton University Press.

Inglehart, R. (1981) Post-materialism in an environment of insecurity. *The American Political Science Review* 75:880-900.

Inglehart, R. (1985) Aggregate stability and individual-level flux in mass belief systems: The level of analysis paradox. *The American Political Science Review* 79:97-116.

Inglehart, R. (1990) *Culture Shift in Advanced Industrial Society*. Princeton: Princeton University Press.

Inglehart, R. (1997) *Modernization and Postmodernization. Cultural, Economic, and Political Change in 43 Societies.* Princeton: Princeton University Press.

Jagodzinski, W. & Manabe, K. (2009) On the similarity of religiosity in different cultures. In: Haller, M., Jowell, R. and Smith, T. W. (eds.), The International Social Survey Programme, 1984-2009. New York: Routledge, 313-336.

Jagodzinski, W. (1983) Materialism in Japan reconsidered: Toward a synthesis of generational and life-cycle explanations. *American Political Science Review* 77:887-894.

Jagodzinski, W. (2004) Methodological problems of value research. In: Vinken, H., Soeter, J. and Ester, P. (eds.) *Comparing Cultures. Dimensions of Culture in a Comparative Perspective*. Leiden & Boston: Brill, 97-121.

Jagodzinski, W. (2010a) Economic, social, and cultural determinants of life satisfaction: Are there differences between Asia and Europe?. *Social Indicators Research* 97:85-104.

Jagodzinski, W. (2010b) Is there a Conflict between Internal and External Validity Or: How Valid are Secular-Rational and Self-Expression Values? Paper presented at the Second European Value Studies Workshop after the 2008 Wave, University of Zagreb, June 7-11, 2010, retrieved from: http://info1.gesis.org/evs2008/Presentation [22.05.2012].

Jenkins, R. (1996) *Social identity*. London: Routledge.

Jensen, J. P. & Bergin, A. E. (1988) Mental health values of professional therapists: A national interdisciplinary survey. *Professional Psychology: Research and Practice* 19:290-297.

Joas, H. & Wiegart, K. (eds.) (2006) Die kulturellen Werte Europas. Frankfurt a. M.: Fischer.

Jöreskog, K. G. (1971) Simultaneous factor analysis in several populations. *Psychometrika* 36:409-426.

Jöreskog, K. G., & Sörbom, D. (1996) *LISREL 8 user's reference guide*. Chicago: Scientific Software International.

Kane, M. T. (2006) Validation. In: Brennan, R. L. (ed) *Educational Measurement (4th Edition)*. Washington, DC: American Council on Education, 17–64.

Kluckhohn, C. (1951) Values and value-orientations in the theory of action: An exploration in definition and classification. In: Parsons, T. & Shils, E. (eds.) *Towards a General Theory of Action*. Cambridge (MA): Harvard University Press, 388-433.

Knoppen, D. & Saris, W. (2009a) Do we have to combine values in the Schwartz' Human Values Scale? A Comment on the Davidov studies. *Survey Research Methods* 3:91-103.

Knoppen, D. & Saris, W. (2009b) Schwartz's Theory of Human Values: Balancing Homogeneity of Reflective Items and Theoretical Coverage. RECSM Working Paper Number 9. Barcelona.

Kosterman, R. & Feshbach, S. (1989) Toward a measure of patriotic and nationalistic attitudes. *Political Psychology* 10:257–274.

Laffan, B. (2004) The European Union and Its Institutions as 'Identity Builders'. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 75–96.

Latcheva, R. 2009. Cognitive Interviewing and Factor-Analytic techniques: A Mixed Method Approach to Validity of Survey Questions Measuring National Identity. *Quality & Quantity* 45:1175-1199.

Lee, N., John W. C., & Chamberlain, L. (2013) The MIMIC model and formative variables: problems and solutions. *AMS Review* 3:3–17.

Lee, S., & Hershberger, S. L. (1990) A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research* 25:313–334.

Licata, L., & Klein, O. (2002) Does European citizenship breed xenophobia? European identification as a predictor of intolerance towards immigrants. *Journal of Community and Applied Social Psychology* 12:1-15.

Little, T. D., Slegers, D. W., & Card, N. A. (2006) A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling*13:59-72.

Lubke, G. H. & Muthén, B. O. (2004) Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling* 11: 514-534.

Lumsden, J. (1980) Variations on a theme by Thurstone. *Applied Psychological Measurement* 4:1-7.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993) The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin* 114:185–199.

Magidson, J. & Vermunt, J. K. (2001) Latent class factor and cluster models, bi-Plots, and related graphical displays. *Sociological Methodology*. 31:223–264.

Maraun, M. D. (2003) *Myths and confusions: Psychometrics and the latent variable model*. Unpublished manuscript, retrieved from: http://www.sfu.ca/~maraun/Mikes%20page%20Myths%20and%20Confusions.html [04.07.2014]

Maraun, M. D.& Gabriel, S. M. (2013) Illegitimate concept equating in the partial fusion of construct validation theory and latent variable modeling. *New Ideas in Psychology* 31:32–42.

Maraun, M.D. & Peters, J. (2005) What Does It Mean That an Issue Is Conceptual in Nature? *Journal of Personality Assessment* 85:128–133.

Margolis E. & Lawrence, S. (2011) Concepts. *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab at Stanford University. Retrieved from: http://plato.stanford.edu/entries/concepts/ [022.08.2014]

Marsh, H. W., Hau, K.-T. & Wen, Z. (2004) In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling* 11:320-341.

Marsh, H.W., Dowson, M.,Pietsch, J., & Walker, R. (2004) Why Multicollinearity matters: A Reexamination of Relations between Self-Efficacy, Self Concept and Achievement. *Journal of Educational Psychology* 3, 518-522.

McGarty, C. (2001) Social Identity Theory does not maintain that identification produces bias, and self-categorization Theory does not maintain that salience is identification: Two comments on Mummendey, Klink and Brown. *British Journal of Social Psychology*, 40:173-176.

Mclaren, L. M. (2006) *Identity, Interests and Attitudes to European Integration*. Basingstoke: Palgrave.

Meredith, W. (1993) Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58:525–543.

Messick, S. (1989) Validity. In: Linn, R. (ed.) *Educational measurement (3rd edition)* Washington, DC: American Council on Education, 13–100.

Meuleman, B., Davidov, E., Schmidt P., & Billiet, J. (2013) Social location and value priorities. A European-wide comparison of the relation between socio-demographic variables and human values. In: Gabriel, O. and Keil, S. I. (eds.) *Society and democracy in Europe*. New York: Routledge, 45-67.

Michell, J. (2000) Normal Science, Pathological Science and Psychometrics. *Theory & Psychology* 10:639–667.

Michell, J. (2008) Is Psychometrics Pathological Science? *Measurement: Interdisciplinary Research & Perspective* 6:7–24.

Michell, J. (2009) Invalidity in Validity. In: Lissitz, R. W (ed.) *The concept of validity. Revisions, new directions, and applications.* Charlotte, NC: Information Age Publishing, 111–133.

Millsap, R. E. & J. Yun-Tein (2004) Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research* 39:479-515.

Mühler, K. & Opp, K. (2004) *Region und Nation. Zu den Ursachen und Wirkungen regionaler und überregionaler Identifikation.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Mummendey, A. & Waldzus, S. (2004) National Difference and European Plurality: Discrimination or Tolerance between European Countries. In: Herrmann, R. K., Risse, T. & Brewer, Marilyn B. (eds.) Transnational Identities. Becoming European in the EU. Lanham: Rowmann & Littlefield, 59-72.

Muthén, B. O. & Asparouhov, T. (2002) *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus.* Mplus Web Notes, No. 4, retrieved from: http://www.statmodel.com/examples/webnote.shtml [23.04.2009]

Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997) Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Technical report, retrieved from: http://www.gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf [18.04.2009]

Muthén, L. K & Muthén, B. O. (1998) *Mplus user's guide*. Los Angeles: Muthén & Muthén.

Muthén, L. K. & Muthén, B. O. (2007) *Mplus user's guid*e. Los Angeles: Muthén & Muthén.

Newton, P. E. (2012) Clarifying the Consensus Definition of Validity. *Measurement: Interdisciplinary Research & Perspective* 10:1–29.

Norris, P. & Inglehart, R. (2004). *Sacred and Secular. Religion and Politics Worldwide*. Cambridge: Cambridge University Press.

Oberski, D. L. (2014) Evaluating Sensitivity of Parameters of Interest to Measurement Invariance in Latent Variable Models. *Political Analysis* 22:45–60.

Opp, K. (2003) Die Identifikation der Bürger mit Europa. Einige Hypothesen und ihre empirische Überprüfung. In: Allmendinger, J. (ed.) *Entstaatlichung und soziale Sicherheit.* Opladen: Leske + Budrich, 475–497.

Opp, K. D. and Wippler, R. (Eds.) (1990) Empirischer Theorienvergleich. Erklärungen sozialen Verhaltens in Problemsituationen. Westdeutscher Verlag, Opladen.

Parsons, T. (1991) *The social system*. London: Routledge.

Perrinjaquet, A., Furrer, O., Usunier, JC.,Cestre, G. & Valette-Florence, P. (2007) A test of the quasi-circumplex structure of human values. *Journal of Research in Personality* 41:820–840.

Pichler, F. (2008) How Real is cosmopolitanism in Europe? *Sociology* 44:1107–1126.

Piurko, Y., Schwartz, S. H., & Davidov, E. (2011) Basic personal values and the meaning of left-right political orientations in 20 countries. *Political Psychology* 32:537-561.

Popper, K. R. (1959) *The logic of scientific discovery*. London: Hutchinson & Sons.

Ramaswamy, V., DeSarbo, W., Reibstein, D. & Robinson, W. (1993) An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science* 12:103–124.

Reicher, S. & Hopkins, N. (2001) *Self and Nation*. London: Sage

Reicher, S. (2004) The Context of Social Identity: Domination, Resistance, and Change. *Political Psychology*, 25:921–945.

Rigdon, E. E. (2012) Rethinking Partial Least Squares Path Modeling: In Praise of Simple Methods. *Long Range Planning* 45:341–358.

Rigdon, E. E. (2013) Lee, Cadogan, and Chamberlain: an excellent point . . . But what about that iceberg? *AMS Review* 3:24–29.

Rokeach, M. (1973*) The Nature of Human Values*. New York: The Free Press.

Sagiv, L. & Schwartz, S. H. (1995) Value priorities and readiness for out-group social contact. *Journal of Personality and Social Psychology* 69, 437-448.

Sagiv, L. & Schwartz, S. H. (2000) Value priorities and subjective well-being: Direct relations and congruity effects. *European Journal of Social Psychology* 30:177-198.

Satorra, A. & Bentler, P. M. (1990) Model conditions for asymptotic robustness in the analysis of linear relations. *Computational Statistics and Data Analysis* 10:235–249.

Satorra, A. (1992) Asymptotic robust inferences in the analysis of mean and covariance structures. *Sociological methodology 22:249-278.*

Schatz, R. T., Staub, E. & Lavine, H. (1999) On the varieties of national attachment: Blind versus constructive patriotism. *Political Psychology* 20:151–174.

Schmidt, P., Bamberg, S., Davidov, E., Herrmann J., & Schwartz, S. H. (2007) Die Messung von Werten mit dem 'Portraits Value Questionnaire' [The measurement of

values with the 'Portraits Value Questionnaire']. *Zeitschrift für Sozialpsychologie* 38:261-275.

Schmitt, H. (2003) The Eurobarometers: Their Evolution, Obvious Merits, and Ways to Add Value to them. *European Union Politics* 4:243-251.

Schnell, R., Hill, P. B., and Esser, E. (2008) *Methoden der empirischen Sozialforschung*. München: Oldenbourg.

Schüßler, A. & Schmidt, P. (2008) Does cognitive interviewing matter? Testing the effects of cognitive methods on data quality. Paper presented at the 7th International Conference on Social Sciences Methodology (RC 33, ISA), Neapel, Italy.

Schwartz, S. H. & Huismans, S. (1995) Value priorities and religiosity in four Western religions. *Social Psychology Quarterly* 58:88-107.

Schwartz, S. H. & Melech, G. (2000) National differences in micro and macro worry: Social, economic and cultural explanations. In: Diener, E. & Suh, E. M. (eds.), *Culture and Subjective Well-Being*. Cambridge (MA): MIT Press, 219-256.

Schwartz, S. H. & Rubel, T. (2005) Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology* 89:1010-1028.

Schwartz, S. H. & Rubel-Lifschitz, T. (2009) Cross-national variation in the size of sex differences in values: Effects of gender equality. *Journal of Personality and Social Psychology* 97:171-185.

Schwartz, S. H. & Vecchione, M. (2011) A Revised Theory and New Instrument to Measure Basic Individual Values. Paper presented at the Regional Conference of the International Association of Cross-Cultural Psychology (IACCP), Istanbul, Turkey, June 30 – July 3, 2011.

Schwartz, S. H. (1992) Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology* 25, 1-65.

Schwartz, S. H. (1994) Are there Universal Aspects in the Content and Structure of Values? *Journal of Social Issues* 50:19–45.

Schwartz, S. H. (2005) Basic human values: Their content and structure across countries. In: Tamayo, A. & Porto, J. B. (eds.), Valores e Comportamento nas Organizações [Values and Behavior in Organizations] Petrópolis, Brazil: Vozes, 21-55.

Schwartz, S. H. (2006) A Theory of Cultural Value Orientations: Explication and Applications *Comparative Sociology* 5:137–182.

Schwartz, S. H. (2006) Les valeurs de base de la personne: Théorie, mesures et applications [Basic human values: Theory, measurement, and applications]. *Revue Française de Sociologie* 47:929-968.

Schwartz, S. H. (2007) Value orientations: Measurement, antecedents and consequences across nations. In: Jowell, R., Roberts, C., Fitzgerald, R., & Eva, G. (eds.), *Measuring Attitudes Cross-Nationally. Lessons from the European Social Survey*. London: Sage, 169-203.

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O. & Konty,

M. (2012) Refining the theory of basic individual values. *Journal of Personality and Social Psychology* 103:663-688.

Schwartz, S. H., Melech, G., Lehmann A., Burgess, S., Harris, M., & Owens, V. (2001) Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross Cultural Psychology* 32:519-542.

Schwartz, S. H., Sagiv, L., & Boehnke, K. (2000) Worries and values. *Journal of Personality* 68:309-346.

Sireci, S. G. (2009) Packing and unpacking sources of validity evidence: history repeats itself again. : Lissitz, R. W (ed.) *The concept of validity. Revisions, new directions, and applications.* Charlotte, NC: Information Age Publishing, 19–37.

Skrobanek, J. (2004) Soziale Identität und Ausländerfeindlichkeit. *Berliner Journal für Soziologie* 14: 357–377.

Skrondal, A. & Rabe-Hesketh, S. (2004) *Generalized latent variable modeling. Multilevel, longitudinal, and structural equation models*. Boca Raton: Chapman & Hall/CRC.

Sörbom, D. (1974) A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology* 27:229-239.

Sörbom, D. (1978) An alternative to the methodology for analysis of covariance. *Psychometrika* 43:381-396.

Steenkamp, J.-B. E. M. & Baumgartner, H. (1998) Assessing measurement invariance in cross national consumer research. *Journal of Consumer Research* 25:78-90.

Steinmetz, H., Isidor, R. & Bäuerle, N. (2012) Testing the Circular Structure of Human Values: A Meta-Analytical Structural Equation Modelling Approach. *Survey Research Methods* 6: 61-75.

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009) Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Quality & Quantity* 43:599-616.

Stevens, S. S. (1946) On the theory of scales of measurement. *Science* 103: 677–680.

Strupp, H. H. (1980) Humanism and psychotherapy: A personal statement of the therapist's essential values. *Psychotherapy* 17:396-400.

Taifel, H. (1978) *Differentiation between Social Groups*. London: Academic Press.

Tajfel, H. & Turner, J. C. (1986) The social identity theory of intergroup behavior. In: Austin, W. G. & Worchel, S. (eds.) *Psychology of Intergroup Relations.* Chicago: Nelson-Hall Publishers, 7–24.

Temme, D. (2006) Assessing measurement invariance of ordinal indicators in cross-national research. In Diehl, S. & Terlutter, R. (eds.*) International advertising and communication. Current insights and empirical findings*. Wiesbaden, Germany: GWV, 455- 472.

Tilly, C (2003) Political identities in changing polities. *Social Research* 70:605–620.

Tilly, C. (2002) *Stories, Identities, and Political Change*. Lanham: Rowman & Littlefield.

Tourangeau, R. (1984) Cognitive science and survey methods. In: Jabine, T. B., Straf, M. L., Tanur, J. M. & Tourangeau, R. (eds.) *Cognitive Aspects of Survey Methodology. Building a Bridge Between Disciplines*. Washington: National Academic Press, 73-100.

van de Vijver, F. J. R. and Leung, K. (1997) *Methods and data analysis for cross-cultural research.* Thousand Oaks: Sage Publications.

Vandenberg, R. J. & Lance, C. E. (2000) A review and synthesis of the measurement invariance literature: Suggestions, practices and recommendations for organizational research. *Organizational Research Methods* 3:4-69.

Vandenberg, R. J. (2002) Towards a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods* 5:139-158.

Verkasalo, M., Lönquist, J. E., Lipsanen J. & Helkama, K. (2009) European Norms and Equations for a two dimensional presentation of values as measured with Schwartz 21 item portrait values questionnaire. *European Journal of Social Psychology* 39:780-792.

Veroff, J., Reuman, D., & Feld, S. (1984) Motives in American men and women across the adult life span. *Developmental Psychology* 20:1142-1158.

Waldzus, S. & Mummendey, A. (2004) Inclusion in a superordinate category, in-group prototypicality, and attitudes towards out-groups. *Journal of Experimental Social Psychology* 40: 466–477.

Waldzus, S., Mummendey, A., Wenzel, M. & Weber, U. (2003) Towards tolerance: Representations of superordinate categories and perceived ingroup prototypicality. *Journal of Experimental Social Psychology* 39:31–47.

Wallström, M. (2006) Understanding public opinion. Madrid: Opening speech at the stakeholders' conference on "Understanding public opinion", 27 October 2006.

Weiler, J. H. H. (1999) *The Constitution of Europe. Essays on the Ends and Means of European Integration.* Cambridge: Cambridge University Press.

Welkenhuysen-Gybels, J. (2004) The performance of some observed and unobserved conditional invariance techniques for the detection of differential item functioning. Quality *& Quantity* 38:681-702.

Welkenhuysen-Gybels, J., & Billiet, J. (2002) A comparison of techniques for detecting cross-cultural inequivalence at the item level. *Quality & Quantity* 36:197-218.

Welzel, C. (2010) How selfish are self-expression values? A civicness test. *Journal of Cross-Cultural Psychology* 41:152-174.

Willis, G. B. (2005) *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks: Sage.

Willis, G. B. 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. London: Sage.

Wimmer, A. (2002) *Nationalist Exclusion and Ethnic Conflict. Shadows of Modernity.* Cambridge: Cambridge University Press.

World Value Survey (2005) official data file v.20090901, 2009. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS, Madrid.

Yack, B (1998) Can patriotism save us from nationalism? Rejoinder to Viroli. *Critical Review* 12:203–206.

Zumbo, B. D. (2009) Validity as contextualized and pragmatic explanation, and its implications for validation practice. In: Lissitz, R. W. (ed.),The concept of validity: Revisions, new directions, and applications. Charlotte, NC: Information Age Publishing, 65–82.

Curriculum Vitae

Georg Datler

Date of birth	14.06.1980
Citizenship	Austrian
Contact	Austrasse 43 8045 Zürich georg.datler@gmx.at +41 (0)76 7416395

Education

09/2009-11/2015	PhD Studies, Sociology, University of Zurich
09/1998-11/2008	Masters degree, Sociology, University of Vienna

Work, academic

09/2009-01/2015	University of Zurich, Institute for Sociology (Prof. Eldad Davidov)
03/2008-02/2009	University of Vienna, Institute for Sociology, project assistant, "Social Dynamics in Vienna" (lead: Prof. Christoph Reinprecht)
10/2007-03/2008	University of Vienna, Institute for Sociology, freelancer, methodological and statistical coaching for a research project on volunteer organisations of medical patients and relatives of medical patients (lead: Prof. Rudolf Forster)
09/2005-06/2007	University of Vienna, Institute for Sociology, Student assistant (Prof. Christoph Reinprecht, Prof. Göran Therborn):
10/2004-01/2006	Austrian Institute for Youth Research, in cooperation with Slovak Academy of Sciences, project collaborator, project „New Communities – New Identities?" (lead: Reingard Spannring)
03/2006-07/2006	Austrian Statistical Office, freelancer, Analysis on differential response rates and panel mortality in EU-SILC
05/2003-12/2005	Institute for Advanced Studies, Vienna, Scholarship for the 5th Framework project „Young Men and Women and their Attitudes towards Citizenship and European Identity" (lead: Prof. Claire Wallace, Prof. Lynn Jamieson)
06/2005-09/2005	University of Vienna, Institute for Sociology, Austrian Social Survey, Religion in Austria (lead Prof. Wolfgang Schulz)

Work, other

07/2007-06/2009	Austrian Statistical Office, project EU-SILC
10/2001-12/2007	social worker in a drug counselling office in Vienna, 1–2 day per week
07/1999-12/2002	pet shop, aquarium division, 1–2 days per week
10/2000-09/2001	alternative civilian service in a drug counselling office in Vienna

Academic teaching

08/2014	Swiss Summer School, course “Structural Equation Modeling” (together with Prof. Eldad Davidov)
fall 2014	Sociology of European Integration, “Die Gesellschaften Europas oder die Europäische Gesellschaft”, University of Zurich, BA Sociology
2010-2014 (spring terms)	Methods lab I, University of Zurich, BA Sociology
fall 2013	Advanced quantitative methods, University of Zurich, MA Sociology (together with Vanita Matta, Daniel Seddig)
06/2010	Course “Structural Equation Modeling with Mplus” (together with Prof. Peter Schmidt), University of Ljubljana
02/2010	Workshop “Structural Equation Modeling with AMOS”, University of Konstanz, Department of Political Science
11/2009	Course “Structural Equation Modeling with MPLUS” (together with Prof. Peter Schmidt, Prof. Eldad Davidov, Bart Meuleman), GESIS-Mannheim (ZUMA)
09/2007-06/2009	Urban sociology, University of Vienna (together with Prof. Christoph Reinprecht)
08/2009	Teaching assistant, ECPR Summer School in Methods and Techniques, Ljubljana, SEM course, Prof. Peter Schmidt, Prof. Eldad Davidov
11/2009	Teaching assistant, 5th ESS Train on “Structural Equation Modeling for Cross-Cultural Research with the Program AMOS”, Mannheim, Prof. Peter Schmidt, Prof. Eldad Davidov
08/2005, 08/2006, 07/2008, 07/2009	Teaching Fellow, Essex Summer School in Social Science Data Analysis and Collection, Advanced Course in SEM (Longitudinal Data, Cross-cultural comparisons), Prof. Peter Schmidt, Prof. Eldad Davidov
fall 2007	Introduction to sociology, University of Vienna
09/2004-06/2005	Student assistant, course on data collection and data analysis, Prof. Wolfgang Schulz

Advanced methods and theory courses/workshops (participant)

06/2010	QMSS II Seminar, “Multilevel Analysis and Multigroup Comparisons in Cross-national Research”, Leuven
011/2009	Problems of Rational Choice Theory, Prof. Siegwart Lindenberg, ETH Zurich
06/2009	QMSS II Seminar, “Quality and Comparability of Measures for Constructs in Comparative Research: Methods and Applications”, Bozen,
05/2004	Multilevel Analysis, Prof. Tom Snijders, Institute for Advanced Studies, Vienna

Languages

German	first language
English	good speaking, reading, and writing
French	basic speaking, fair reading, basic writing
Spanish	basic speaking and reading

Software

Advanced	MS Office, AMOS, MPLUS, SPSS
Basic	Atlas.ti, LISREL, R, SAS, Stata,

Publications

Journal articles and book chapters

Datler, Georg (forthcoming) European identity as a safeguard against xenophobia? A differentiated view based on identity content. In: Grimm, Jürgen, Huddy, Leonie, Schmidt, Peter, and Seethaler, Josef (eds.) Dynamics of National Identity. London: Routledge.

Schroedter, Julia H.; Rössel, Jörg; Datler, Georg (2015): European Identity in Switzerland: The Role of Intermarriage, Transnational Social Relations and Experiences. The Annals of the American Academy of Political and Social Science 662 (1): 148-168.

Datler, Georg, Jagodzinski, Wolfgang, Schmidt, Peter (2013) Two Theories on the Test Bench: Internal and External Validity of the Theories of Ronald Inglehart and Shalom Schwartz. Social Science Research 42 (3), 906–925.

Das Konzept der Europäischen Identität jenseits der Demos-Fiktion (2012) Aus Politik und Zeitgeschichte 4/2012, 57–61.
[*The concept of European identity beyond the fiction of a demos*]

Latcheva, Rossalina, Datler, Georg, Rossbacher, Eva (2012) The Concept of European Identity: Overused and Underspecified? In: Salzborn, Samuel, Davidov, Eldad, Reinecke, Jost (eds.) Methods, theories, and empirical applications in the social sciences. Festschrift for Peter Schmidt, Wiesbaden: VS Verlag, 235–248.

Davidov, Eldad, Georg Datler, Peter Schmidt, and Schwartz, Shalom H (2011) Testing the invariance of values in the Benelux countries with the European Social Survey: Accounting for ordinality. In: Davidov, Eldad, Peter Schmidt and Jacques Billiet (eds.) Methods and Applications in Cross-Cultural Analysis. NJ: Taylor and Francis, 149–168.

Datler, Georg (2010) "Europäische Identität" in der empirischen Sozialforschung: Ein Konzept ohne Inhalt? Forum 21, European Journal on Child and Youth Research 6, 68-75.
[*"European Identity" in empirical social research: A concept without content?*]

Datler, Georg, Mahidi, Margarethe (2009) Armutsgefährdung und Wohnsituation. Statistische Nachrichten 6/2009, 458–473.
[*Risk of poverty and housing*]

Wallace, Claire, Reingard Spannring, and Datler, Georg (2008) What Leads Young People to Identify with Europe? An Exploration of the Impact of Exposure to Europe and Political Engagement on European Identity among Young Europeans (2008). Perspectives on European Politics and Society 9 (4), 480–498.

Datler, Georg, Kerschbaum, Johann, and Schulz, Wolfgang. Religion und Kirche in Österreich—Bekenntnis ohne Folgen? (2005). SWS-Rundschau, 4/2005.
[*Religion and church attendance in Austria – Denomination without commitment?*]

Datler, Georg, Wallace, Claire, and Spannring, Reingard (2005). What leads young people to identify with Europe? IHS Sociological Series 69.

Wallace, Claire, Datler, Georg and, Spannring, Reingard.Young People and European Citizenship (2005) IHS Sociological Series 68.

Spannring, Reingard, Wallace, Claire, and Datler, Georg (2004) If you have a grandpa, send him to Europe. Attitudes of young Austrians towards the EU elections. Sociológia 36 (3), 253-272.

Project reports (selection)

Reinprecht, Christoph, Georg Datler, Carmen Keckeis, and Kurtev, Angelina (2009) Soziale Dynamik im Stadtraum. Sozialraumanalysen in 8 Wiener Wohnvierteln. Muncipality of Vienna, Vienna.
[*Social dynamics in Vienna. Social area analysis in 8 neighbourhoods*]

Dalter, Georg, Matthias Till (2009) Armutslagen und Chancen für Eingliederung in Österreich. Indikatorenliste mit aktualisierten Daten. Statistik Austria, Vienna.
[*Poverty and potentials for integration in Austria. Set of indicators with new data*]

Till, Matthias, Eiffe, Franz F. , Datler, Georg, Henke, Justus, Schrittwieser, Karin, Till-Tentschert, Ursula and Wagner-Pinter, Michael (2009) Armutslagen und Chancen für Eingliederung in Österreich. Indikatoren für das Monitoring des nationalen Strategieplans 2008–11. Statistik Austria, Vienna.
[*Poverty and potentials for social inclusion in Austria. Indicators for the montoring of the national strategic agenda 2008–11*]

Datler, Georg, Till, Matthias, and Skina, Magdelena.EU-SILC Sonderauswertung zur sozialen Lage älterer Menschen. Modul 1: Ausgangslange und tabellarischer Arbeitsbehelf (2008) Statistik Austria, Vienna.
[EU-SILC analyses of the living conditions of elderly people, part 1: Basic figures]

Datler, Georg (2006) EU-SILC 2005/2006: Ausschöpfung im Panel (2006) Statistik Austria, Vienna.
[EU-SILC 2005/2006: Panel Attrition]

Spannring, Reingard, Lasticova, Barbara, Waechter, Natalie, Datler, Georg, Petrjánošová, Magda, and Bianchi, Gabriel (2005) Neue Gemeinschaften—Neue Identitäten? Eine Studie zu territorialen Identitäten Jugendlicher in der österreichisch-slowakischen Grenzregion. Endbericht OeNB Jubiläumsfonds Projektnr: 10690.
[*New Communities—New Identities? **A study of territorial identities of young people at the Austrian-Slovak border***]

Conference Presentations (selection)

European identity. A variable Variable (2014) University of Bern, Institute of Sociology, Research Colloquium.

What is Europe? Meanings of Europe in different social contexts (2014) Council of European Studies, 21st International Conference of Europeanists, Washington. (together with Jörg Rössel and Julia Schroedter)

European identity as a safeguard against xenophobia? A differentiated view based on identity content (2011) 1st Vienna Forum of Social Sciences: Identity - Diversity - Integration, Vienna.

European, national, and regional identities: An empirical analysis of their interrelations and dynamics (2011). ESA conference, Geneva.

The concept of European identity: empirical social science as the puppet of normative claims? (2011) ISPP Annual Scientic Meeting, Istanbul.

Measuring European Identity: A Lack of Content? (2010) ISA congress, Gotenburg.

What leads young people to identify with Europe? An exploration of “exposure” to Europe and resources for identification among young Europeans (2005) ESA conference, Torun.

Research Briefing 5: Active Citizens? The engagement of young people in Europe (2004) Final Conference of the Project "Orientations of Young Men and Women to Citizenship and European Identity", Brussels. (together with Claire Wallace)